

UNIVERSIDAD COMPLUTENSE DE MADRID  
FACULTAD DE CIENCIAS BIOLÓGICAS  
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR



BIOLOGÍA ESTRUCTURAL DE INOSITOL KINASAS, ENZIMAS  
IMPLICADAS EN LA REGULACIÓN DE INOSITOL FOSFATOS

TESIS DOCTORAL DE:  
**JOSÉ IGNACIO BAÑOS SANZ**

BAJO LA DIRECCIÓN DE:  
**BEATRIZ GONZÁLEZ PÉREZ**  
**JULIANA SANZ APARICIO**

Madrid, 2013

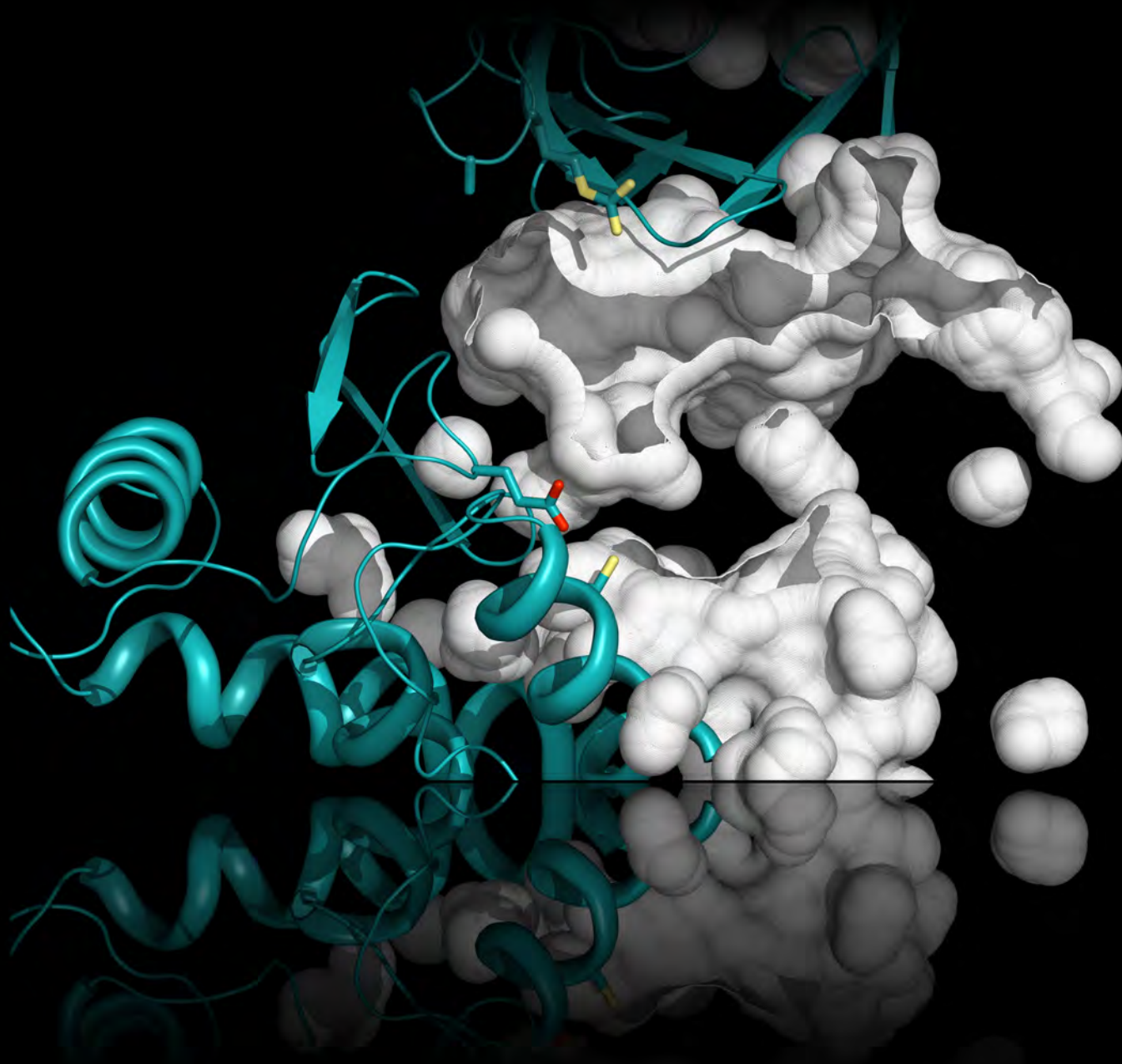
UNIVERSIDAD COMPLUTENSE DE MADRID  
FACULTAD DE CIENCIAS BIOLÓGICAS  
DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I

---

# BIOLOGÍA ESTRUCTURAL DE INOSITOL KINASAS,

enzimas implicadas en la regulación de inositol fosfatos

---



---

TESIS DOCTORAL  
JOSÉ IGNACIO BAÑOS SANZ  
2013



**DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA  
MOLECULAR I  
FACULTAD DE CIENCIAS BIOLÓGICAS  
UNIVERSIDAD COMPLUTENSE DE MADRID**

**Biología estructural de inositol kinasas, enzimas  
implicadas en la regulación de inositol fosfatos.**



**Memoria presentada por José Ignacio Baños Sanz  
para optar al grado de Doctor.**

Esta Tesis Doctoral ha sido realizada en el Departamento de Cristalografía y Biología Estructural del Instituto de Química Física "Rocasolano" (Consejo Superior de Investigaciones Científicas, CSIC) bajo la dirección de la Dra. Beatriz González Pérez (Científico Titular) y la Dra Juliana Sanz Aparicio (Investigador Científico) ambas miembros del Departamento de Cristalografía y Biología Estructural del Instituto de Química Física "Rocasolano" (Consejo Superior de Investigaciones Científicas, CSIC)

**MADRID, 2013**





La Dra. Beatriz González Pérez, científico titular del Departamento de Cristalografía y Biología Estructural del Instituto de Química Física “Rocasolano” (Consejo Superior de Investigaciones Científicas, CSIC) y la Dra. Juliana Sanz Aparicio, investigador científico del Departamento de Cristalografía y Biología Estructural del Instituto de Química Física “Rocasolano” (Consejo Superior de Investigaciones Científicas, CSIC) , certifican que:

Don José Ignacio Baños Sanz, Licenciado en Biología por la Universidad Complutense de Madrid, ha realizado bajo nuestra dirección el trabajo titulado “Biología Estructural de Inositol Kinasas, enzimas implicadas en la regulación de inositol fosfatos”, en el Departamento de Cristalografía y Biología Estructural del Instituto de Química Física “Rocasolano” (Consejo Superior de Investigaciones Científicas, CSIC).

En nuestra opinión, este trabajo reúne las condiciones exigidas por la legislación vigente y tiene la originalidad, el rigor y la calidad científica necesarios y suficientes para ser presentado por Don José Ignacio Baños Sanz para optar al Grado de Doctor en Biología por la Universidad Complutense de Madrid.

Para que así conste y con los efectos oportunos, firmamos el presente certificado.

En Madrid a 14 de Junio de 2013

Dra. Beatriz González Pérez  
Científico titular CSIC

Dra. Juliana Sanz Aparicio  
Investigador científico CSIC



Esta tesis fue realizada gracias a la concesión a José Ignacio Baños Sanz de una beca predoctoral de Formación de Profesorado Universitario (FPU) (AP2008-00916) durante el periodo comprendido entre Julio de 2009 y Julio 2013 por parte del Ministerio de Educación.

The Spanish Ministry of Education awarded José Ignacio Baños Sanz with an Formación de Profesorado Universitario (FPU) fellowship (AP2008-00916) from July 2009 to July, 2013 to develop this thesis.

La Federación Europea de Sociedades de Bioquímica (FEBS) concedió a José Ignacio Baños Sanz una Short-Term Fellowship para realizar una estancia corta en el Grupo “Synchrotron Crystallography” del laboratorio Europeo de Biología Molecular situado en Grenoble (Francia) entre Octubre-Noviembre de 2012.

The Federation of European Biochemical Societies (FEBS) awarded José Ignacio Baños Sanz with a Short-Term Fellowship to perform a short research stay in the Synchrotron Crystallography Team (European Molecular Biology Laboratory, Grenoble, France) from October to November, 2012.



*A mis padres  
y mi hermana*



# Agradecimientos

Querría agradecer a la Dra. Beatriz González la dirección y a la Dra. Juliana Sanz-Aparicio la codirección de esta tesis.

A todos los investigadores principales del Departamento de Cristalografía y Biología Estructural del IQFR, así como a todos los compañeros del Departamento su apoyo durante este tiempo. A nivel “técnico” querría agradecerle a la Dra. Juana María González Rubio su gran trabajo en el laboratorio de biología molecular y a Maider Villate su enorme trabajo en la cristalización de algunos de los complejos de IP<sub>5</sub> 2-K nativa presentados en esta tesis. Por otra parte, querría agradecer al Dr. José Miguel Mancheño Gómez el permitirnos usar su vector pKLSL tan intensivamente, a la Dra. Margarita Menéndez su ayuda en la interpretación de algunos experimentos, y a Javier Gutiérrez su inestimable ayuda con las imágenes de portada y contraportada de esta tesis.

Mucha gente del fuera del IQFR ha colaborado para que este trabajo llegase a buen puerto. En primer lugar, vaya un agradecimiento para el Dr. Charles A. Brearley, que ha colaborado intensamente desde Norwich con buena parte de los ensayos bioquímicos de esta tesis. Querría agradecerle también al Dr. Adam Round y el resto de personal del grupo de “Synchrotron Crystallography” el excelente trato que recibí durante mi estancia en el EMBL de Grenoble, y a la Dra. Inés G. Muñoz (CNIO) que corrigió la parte escrita de SAXS de esta tesis. A los sincrotrones empleados durante esta tesis ESRF, SLS y ALBA y a los respectivos “*local contacts*” (y muy especialmente al Dr. Jordi Juanhuix, del sincrotrón ALBA) que trabajan en dichos sincrotrones para que todo funcione a la perfección. Gracias también a la Dra. Olga Perisic (LMB) por permitirnos usar los vectores allí desarrollados.

A mis amigos (los de dentro y los de fuera de Rocasolano) por estar ahí y darme los empujoncitos en los momentos clave de esta “maratón”.

Por último, querría agradecer a mi familia la descomunal paciencia durante todos estos años de licenciatura+doctorado, y muy especialmente a mis padres y a mi hermana; sin su apoyo esto no hubiera sido posible.

Y a Carmela, por apoyarme día tras día.





*The wonderful thing about science is that it's alive.*

*Richard P. Feynman*

*No coffee, no structure*

*R. Kahn*



# Resumen

Los inositol fosfatos (IPs) son un amplio grupo de moléculas implicadas en señalización celular. La señalización mediante IPs comienza al romperse la cadena del inositol lipídico fosfatidilinositol (1,4,5) trisfosfato (PIP<sub>3</sub>) mediante la acción de la fosfolipasa C, generándose I(1,4,5)P<sub>3</sub> o IP<sub>3</sub>. En concreto, el IP<sub>3</sub> es el más estudiado debido al papel que desempeña como segundo mensajero en la movilización de Ca<sup>2+</sup>, pero otros IPs también presentan papeles relevantes. Otro de los IPs más importantes es el fitato o inositol hexakisfosfato (IP<sub>6</sub>). Entre las funciones del IP<sub>6</sub> se encuentran la regulación de procesos celulares clave tales como la exportación del mRNA, reparación de DNA, apoptosis, y regulación de la estructura de la cromatina. Los niveles de los IPs se encuentran estrictamente regulados por la acción de diferentes kinasas y fosfatas. El presente proyecto de tesis se centra en la biología estructural de las inositol fosfato kinasas (IPKs) que fosforilan las diferentes posiciones del anillo de inositol. Existen varias familias de IPKs, las cuales presentan un dominio catalítico con actividad kinasa, y en numerosas ocasiones otros dominios principalmente reguladores.

En este trabajo se planteó el estudio de dos aspectos clave en el metabolismo de los IPs: la regulación por Ca<sup>2+</sup>/calmodulina de la IP<sub>3</sub> Kinasa (IP<sub>3</sub> 3-K) desde un punto de vista estructural, (inicio en la síntesis de IPs a partir de IP<sub>3</sub>) y la elucidación de la estructura de la primera estructura de una IP<sub>5</sub> 2-K (enzima encargada de la síntesis de IP<sub>6</sub>). Para ello se ha planteado la purificación, cristalización y resolución estructural de la inositol pentakisfosfato 2 kinasa (IP<sub>5</sub> 2-K) de *Arabidopsis thaliana* y la purificación, cristalización y resolución estructural del complejo entre el dominio de unión a CaM (*Calmodulin Binding Domain*, CBD) de la IP<sub>3</sub> 3-K A humana y la calmodulina humana.

La estructura de la IP<sub>5</sub> 2-K de *A. thaliana* presenta un plegamiento de tipo  $\alpha+\beta$ , estando dividida en dos lóbulos, N-terminal y C-terminal. Una gran parte del dominio C-terminal se encuentra formando una región estructural nueva (C<sub>IP</sub>), implicada en la unión al inositol. Además, la IP<sub>5</sub> 2-K presenta un sitio de unión a Zn<sup>2+</sup>. Se han identificado residuos claves para la unión del sustrato y la catálisis, y se ha confirmado su papel mediante mutagénesis dirigida. La cristalización y resolución de las formas libres y unidas a nucleótidos de IP<sub>5</sub> 2-K permitieron a su vez la caracterización de los cambios conformacionales que sufre esta enzima durante su ciclo catalítico.

Por otro lado, la obtención de la estructura del complejo entre el CBD (*Calmodulin Binding Domain*) de la IP<sub>3</sub> 3-K A y CaM ha permitido identificar una nueva clase de motivos de unión a esta proteína. Gracias a la combinación de datos a alta y baja resolución se ha logrado caracterizar el modo de activación de la IP<sub>3</sub> 3-K por la CaM.

En conjunto estos resultados han aportado datos fundamentales que permiten obtener una visión más detallada de la síntesis y regulación de los IPs.

# Índice de Contenidos

<b>ABSTRACT</b>	<b>XV</b>
-----------------	-----------

<b>Structural biology of inositol kinases, enzymes in charge of regulation of inositol phosphates.</b>	<b>xvii</b>
--	-------------

<b>Introduction</b>	<b>xvii</b>
---------------------	-------------

<b>Objetives</b>	<b>xviii</b>
------------------	--------------

<b>Metodology. Materials and Methods</b>	<b>xix</b>
--	------------

<b>Results</b>	<b>xx</b>
----------------	-----------

IP <sub>5</sub> 2-K: complexes with substrates (IP <sub>5</sub> IP <sub>5</sub> +AMPPNP and IP <sub>4</sub> +AMPPNP) and products (IP <sub>6</sub> and IP <sub>6</sub> +AMPPNP).	xx
--	----

IP <sub>5</sub> 2-K: apo form and complexes with nucleotide (AMPPNP and ADP)	xxi
--	-----

IP <sub>3</sub> 3-K regulation by calmodulin	xxi
--	-----

<b>Conclusions</b>	<b>xxii</b>
--------------------	-------------

IP <sub>5</sub> 2-K	xxii
---------------------	------

IP <sub>3</sub> 3-K	xxiii
---------------------	-------

<b>Bibliography</b>	<b>xxiv</b>
---------------------	-------------

<b>1 INTRODUCCIÓN</b>	<b>1</b>
-----------------------	----------

<b>1.1 Los inositidos</b>	<b>3</b>
---------------------------	----------

<b>1.2 Inositoles lipídicos</b>	<b>4</b>
---------------------------------	----------

<b>1.3 Liberación de IP<sub>3</sub> y señalización por Ca<sup>2+</sup></b>	<b>5</b>
--	----------

1.3.1 Introducción	5
--------------------	---

1.3.2 El Ca <sup>2+</sup> como segundo mensajero. Proteínas de unión a Ca <sup>2+</sup>	6
---	---

1.3.3 La calmodulina, un sensor de calcio prototipo	6
---	---

1.3.4 Proteínas reguladas por CaM. Dominios de unión a CaM (CBDs)	7
---	---

<b>1.4 Una nueva vía de señalización celular a partir de IP<sub>3</sub>: la vía de los inositol fosfatos</b>	<b>9</b>
--	----------

1.4.1 Introducción	9
--------------------	---

1.4.2 Inositol fosfato kinasas	10
--------------------------------	----

1.4.2.1 Familia I. Inositol polifosfato kinasas	10
---	----

1.4.2.1.1 Inositol trifosfato 3 kinasas (IP <sub>3</sub> 3-Ks)	11
--	----

1.4.2.1.2 Inositol polifosfato multikinasa	15
--	----

1.4.2.1.3 Inositol hexakisfosfato kinasa	17
--	----

1.4.2.2 Familia II. Inositol (1,3,4) trisfosfato 5/6 kinasas (IP <sub>3</sub> 5/6-Ks)	19
---	----

1.4.2.3 Familia III. Inositol pentakisfosfato 2 kinasas (IP <sub>5</sub> 2-Ks)	20
<b>2 OBJETIVOS</b>	<b>23</b>
<b>3 MATERIALES Y MÉTODOS</b>	<b>27</b>
<b>3.1 Determinación estructural de proteínas mediante cristalografía de rayos X</b>	<b>29</b>
<b>3.2 Metodología experimental</b>	<b>30</b>
3.2.1 IP <sub>5</sub> 2-K de <i>A. thaliana</i>	30
3.2.1.1 Estructura cristalográfica de IP <sub>5</sub> 2-K en complejo con inositidos. Complejos con sustratos (IP <sub>5</sub> , IP <sub>5</sub> /AMPPNP e IP <sub>4</sub> /AMPPNP) y productos (IP <sub>6</sub> e IP <sub>6</sub> /ADP)	30
3.2.1.1.1 Construcción de vectores de expresión	30
3.2.1.1.2 Expresión de proteína para estudios de cristalización	30
3.2.1.1.3 Purificación	31
3.2.1.1.4 Cristalización	31
3.2.1.1.5 Obtención y procesamiento de datos	32
3.2.1.1.5.1 Crioprotección y derivatización con átomos pesados	32
3.2.1.1.5.2 Recogida y procesamiento de datos	33
3.2.1.1.6 Resolución estructural y construcción de los modelos	34
3.2.1.1.6.1 Resolución estructural y refinamiento del complejo IP <sub>5</sub> 2-K IP <sub>6</sub> /ADP/Pb <sup>2+</sup>	34
3.2.1.1.6.2 Resolución estructural y refinamiento de los complejos IP <sub>5</sub> 2-K/IP <sub>6</sub> /ADP, IP <sub>5</sub> 2-K /IP <sub>5</sub> /AMPPNP, IP <sub>5</sub> 2-K /IP <sub>6</sub> , IP <sub>5</sub> 2-K /IP <sub>5</sub> e IP <sub>5</sub> 2-K /IP <sub>4</sub> /AMPPNP	34
3.2.1.1.7 Análisis estructural	35
3.2.1.2 Estructura cristalográfica de IP <sub>5</sub> 2-K en estado libre y en complejo con nucleótido (IP <sub>5</sub> 2-K /AMPPNP e IP <sub>5</sub> 2-K /ADP)	35
3.2.1.2.1 Obtención del vector de expresión para el mutante IP <sub>5</sub> 2-K <sub>W129A</sub>	35
3.2.1.2.2 Expresión y purificación de IP <sub>5</sub> 2-K <sub>W129A</sub> para estudios de cristalización	35
3.2.1.2.3 Cristalización de IP <sub>5</sub> 2-K (IP <sub>5</sub> 2-K <sub>W129A</sub> )	36
3.2.1.2.3.1 IP <sub>5</sub> 2-K <sub>W129A</sub> en estado libre	36
3.2.1.2.3.2 Complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP	36
3.2.1.2.4 Recogida y procesamiento de datos	36
3.2.1.2.5 Resolución estructural y construcción de los modelos	37
3.2.1.2.6 Análisis estructural	37
3.2.1.3 Análisis bioquímico de IP <sub>5</sub> 2-K	38
3.2.1.3.1 Obtención de Mutantes de IP <sub>5</sub> 2-K	38
3.2.1.3.2 Cuantificación de zinc en la muestra	39
3.2.1.3.3 Ensayos cinéticos y de actividad enzimática	39
3.2.1.3.3.1 Ensayos de actividad enzimática de IP <sub>5</sub> 2-K nativa y mutantes	39
3.2.1.3.3.2 Efecto del EDTA sobre la actividad de IP <sub>5</sub> 2-K	40
3.2.1.3.3.3 Ensayos de actividad enzimática de los mutantes W129A, R45A W129A/R130A R130A y ΔS253/S55	40
3.2.1.3.3.4 Ensayos cinéticos de las muestras IP <sub>5</sub> 2-K y IP <sub>5</sub> 2-K <sub>W129A</sub>	40
3.2.1.3.4 Medidas de fluorescencia	40
3.2.1.3.4.1 IP <sub>5</sub> 2-K nativa vs IP <sub>5</sub> 2-K tratada con EDTA	40
3.2.1.3.4.2 Unión de ligandos en IP <sub>5</sub> 2-K nativa vs IP <sub>5</sub> 2-K <sub>W129A</sub>	41
3.2.1.3.5 Ensayos mediante geles nativos	41
3.2.1.3.6 Ensayos de proteólisis limitada	41

3.2.2 Complejo IP <sub>3</sub> 3-K A/CaM de <i>Homo sapiens</i>	42
3.2.2.1 Estructura mediante cristalografía de rayos X del complejo entre el dominio de unión a CaM (CBD) de IP <sub>3</sub> 3-K A y CaM	42
3.2.2.1.1 Construcción de vectores de expresión del CBD de IP <sub>3</sub> 3-K A y CaM humanas	42
3.2.2.1.2 Expresión y purificación de proteína para estudios de cristalización	43
3.2.2.1.2.1 Expresión y purificación de CaM humana	43
3.2.2.1.2.2 Expresión y purificación de la IP <sub>3</sub> - 3K A humana (fragmentos IP <sub>3</sub> 3-K <sub>158-201</sub> , IP <sub>3</sub> 3-K <sub>158-189</sub> e IP <sub>3</sub> 3-K <sub>158-183</sub> )	43
3.2.2.1.2.3 Formación y purificación de los complejos IP <sub>3</sub> 3-K <sub>158-201</sub> /CaM, IP <sub>3</sub> 3-K <sub>158-189</sub> /CaM y IP <sub>3</sub> 3-K <sub>158-183</sub> /CaM	44
3.2.2.1.3 Cristalización del complejo IP <sub>3</sub> 3-K <sub>158-183</sub> /CaM	44
3.2.2.1.4 Recogida y procesado de datos	45
3.2.2.1.5 Resolución estructural y construcción de los modelos	45
3.2.2.1.6 Análisis estructural	46
3.2.2.2 Estructura mediante dispersión a bajo ángulo (SAXS) del complejo IP <sub>3</sub> 3-K/CaM	46
3.2.2.2.1 Expresión y purificación del complejo IP <sub>3</sub> 3-K <sub>(128-461)</sub> para estudios de SAXS	46
3.2.2.2.1.1 Expresión y purificación de CaM	46
3.2.2.2.1.2 Expresión y purificación de IP <sub>3</sub> 3-K <sub>(128-461)</sub>	46
3.2.2.2.1.3 Purificación del complejo IP <sub>3</sub> 3-K <sub>(128-461)</sub> /CaM	47
3.2.2.2.2 Obtención de un modelo del complejo IP <sub>3</sub> 3-K <sub>(128-461)</sub> /CaM mediante SAXS	47
3.2.2.2.2.1 Recogida y procesado de datos de SAXS	47
3.2.2.2.2.2 Obtención de la envoltura y modelado del complejo IP <sub>3</sub> 3-K <sub>(128-461)</sub> /CaM	48

## 4 RESULTADOS Y DISCUSIÓN

51

<b>4.1 Estructura de la IP<sub>5</sub> 2-K de <i>A. thaliana</i>. Complejos con sustratos (IP<sub>5</sub> e IP<sub>5</sub>/AMPPNP) y productos (IP<sub>6</sub> e IP<sub>6</sub>/ADP)</b>	<b>53</b>
4.1.1 Obtención de muestra para estudios de cristalización	53
4.1.2 Cristalización de la IP <sub>5</sub> 2-K y complejos con inositido	54
4.1.3 Resolución estructural de la IP <sub>5</sub> 2-K y construcción de modelos.	56
4.1.4 Estructura de la IP <sub>5</sub> 2-K: descripción general del plegamiento	61
4.1.5 El centro activo de la IP <sub>5</sub> 2-K	64
4.1.5.1 Sitio de unión al nucleótido	64
4.1.5.2 Sitio de unión al inositido	66
4.1.5.3 La unión del metal en el centro activo	68
4.1.5.4 Cambios en el centro activo unidos a la entrada de nucleótido	69
4.1.5.5 El mecanismo catalítico de la IP <sub>5</sub> 2-K	70
4.1.6 El sitio de unión a Zn <sup>2+</sup>	70
4.1.7 La IP <sub>5</sub> 2-K de <i>A. thaliana</i> como modelo para la familia de las IP <sub>5</sub> 2-K	72
4.1.8 La familia de la IP <sub>5</sub> 2-K y su relación con la familia de las IPKs	74
4.1.8.1 La IP <sub>5</sub> 2-K conserva el plegamiento de la familia de IPKs	74
4.1.8.2 Inserciones en las IP <sub>5</sub> 2-K. Comparación con inserciones en las IPKs	75
4.1.8.3 El lóbulo de unión a inositido: un elemento de divergencia entre los diferentes miembros de la familia IPK	77
4.1.8.4 Los residuos catalíticos: conservación a través de la familia IPK	79
4.1.9 La IP <sub>5</sub> 2-K y su relación con las proteínas kinasas (PKs)	79
4.1.10 Análisis mediante mutagénesis dirigida de la IP <sub>5</sub> 2-K	81
4.1.11 Conclusión	82

<b>4.2 Estructura de IP<sub>5</sub> 2-K<sub>W129A</sub> de <i>A. thaliana</i>: apoenzima y complejos binarios con AMPPNP y ADP</b>	<b>84</b>
4.2.1 Introducción	84
4.2.2 Obtención de muestras de IP <sub>5</sub> 2-K <sub>W129A</sub> para estudios de cristalización	85
4.2.3 Cristalización de IP <sub>5</sub> 2-K <sub>W129A</sub>	85
4.2.3.1 Cristalización de la apoenzima	85
4.2.3.2 Cristalización de los complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP	86
4.2.4 Procesado de datos, resolución estructural y construcción de los modelos IP <sub>5</sub> 2-K <sub>W129A</sub> , IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP	86
4.2.5 Estructura de la IP <sub>5</sub> 2-K <sub>W129A</sub>	89
4.2.5.1 Estructura de la IP <sub>5</sub> 2-K <sub>W129A</sub> : conformación abierta	89
4.2.5.2 Sitio de unión al inositido: comparación entre forma abierta y forma cerrada	91
4.2.6 Estructura de los complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP	92
4.2.6.1 Estructura de la IP <sub>5</sub> 2-K <sub>W129A</sub> unida al nucleótido: conformación semicerrada	92
4.2.6.2 Sitio de unión a inositido : comparación entre forma abierta y forma semicerrada	93
4.2.7 Formación de la cremallera (interacción α6-L3) en las diferentes formas de IP <sub>5</sub> 2-K	94
4.2.8 Análisis mediante mutagénesis dirigida y ensayos de actividad	96
4.2.8.1 Análisis mediante mutagénesis dirigida de la cremallera α6-L3 y residuos del N <sub>t</sub> implicados en la unión a IP	96
4.2.8.2 Ensayos cinéticos de IP <sub>5</sub> 2-K e IP <sub>5</sub> 2-K <sub>W129</sub>	97
4.2.9 Papel del lóbulo N <sub>t</sub> en la catálisis enzimática	98
4.2.9.1 Análisis estructural del complejo IP <sub>5</sub> 2-K/I(3,4,5,6)P <sub>4</sub> /AMPPNP	98
4.2.9.2 Discriminación de los diferentes enantiómeros de IP <sub>4</sub> por la IP <sub>5</sub> 2-K	100
4.2.10 Otras pruebas experimentales del cambio conformacional	101
4.2.10.1 Fluorescencia intrínseca de la IP <sub>5</sub> 2-K y la IP <sub>5</sub> 2-K <sub>W129A</sub>	101
4.2.10.2 Ensayos mediante geles nativos	103
4.2.10.3 Ensayos de proteólisis limitada de IP <sub>5</sub> 2-K e IP <sub>5</sub> 2-K <sub>W129A</sub>	104
4.2.11 El ciclo catalítico de la IP <sub>5</sub> 2-K	105
4.2.12 Cambios conformacionales dentro de la familia de las IPKs	106
4.2.13 Conclusión	107
<b>4.3 Regulación de la IP<sub>3</sub> 3-K mediante CaM</b>	<b>108</b>
4.3.1 Obtención de muestra de complejos CaM/CBD IP <sub>3</sub> 3-K	108
4.3.2 Cristalización del complejo CaM/ IP <sub>3</sub> 3-K <sub>(158-183)</sub>	108
4.3.3 Recogida de datos, resolución estructural y refinado del modelo	110
4.3.4 Estructura general del complejo	113
4.3.5 Comparación con otros motivos de unión a CaM (CBDs)	118
4.3.5.1 Comparación con la CaMK I. Motivo tipo 1-10	119
4.3.5.2 Comparación con la sMLCK. Motivo tipo 1-14	122
4.3.5.3 Comparación con la CaMKK. Motivo tipo 1-16	123
4.3.6 Comparación entre las dos formas de CaM observadas en este trabajo	127
4.3.7 Comparación con otras isoformas de la IP <sub>3</sub> 3-K. Correlación con estudios bioquímicos	128
4.3.8 Un modelo del complejo CaM/IP <sub>3</sub> 3-K <sub>(128-461)</sub> mediante la combinación de datos a alta y baja resolución	130
4.3.9 Un posible mecanismo de activación de la IP <sub>3</sub> 3-K mediado por CaM	132
4.3.10 Conclusión	133

## 5 CONCLUSIONES Y PERSPECTIVAS

**135**



<b>5.1 Conclusiones</b>	<b>137</b>
5.1.1 IP <sub>5</sub> 2-K	137
5.1.2 IP <sub>3</sub> 3-K	138
<b>5.2 Perspectivas</b>	<b>138</b>
5.2.1 IP <sub>5</sub> 2-K	138
5.2.2 IP <sub>3</sub> 3-K	139
<b>6 BIBLIOGRAFÍA</b>	<b>141</b>
<b>7 ANEXOS</b>	<b>159</b>
<b>7.1 Anexo 1: resolución de estructuras de proteínas mediante cristalografía de rayos X</b>	<b>161</b>
7.1.1 Obtención de muestras de proteínas	161
7.1.2 Cristalización de proteínas	162
7.1.2.1 Introducción	162
7.1.2.2 Base teórica del proceso de cristalización	163
7.1.2.3 Metodología de cristalización. Búsqueda y mejora de condiciones de cristalización	164
7.1.3 Recogida de datos de difracción	166
7.1.3.1 El fenómeno de la difracción	166
7.1.3.2 Generación de rayos X	167
7.1.3.3 Recogida de datos de difracción	168
7.1.4 Procesado de datos de difracción	169
7.1.5 Resolución estructural	171
7.1.5.1 Introducción: el problema de la fase	171
7.1.5.2 Métodos experimentales de obtención de las fases	172
7.1.5.2.1 Método de dispersión anómala múltiple (MAD)	173
7.1.5.2.2 Método de Reemplazamiento Molecular (MR)	174
7.1.5.3 Mejora de fases	175
7.1.6 Construcción, refinado y validación del modelo	175
7.1.6.1 Construcción y refinado del modelo	175
7.1.6.2 Validación	177
<b>7.2 Anexo 2: materiales empleados en la elaboración de este trabajo</b>	<b>178</b>
7.2.1 Listado de Oligonucleótidos	178
7.2.2 Cepas bacterianas	179
7.2.3 Vectores	180
7.2.4 Reactivos, kits de biología molecular y enzimas	181
7.2.5 Resumen de las construcciones empleadas	182
7.2.6 Resumen de las estructuras resueltas	182
<b>PUBLICACIONES</b>	<b>183</b>

## Índice de figuras

Figura 1.1. Isómeros del inositol.	3
Figura 1.2. Funciones del I(1,4,5)P <sub>3</sub> en la célula.	5
Figura 1.3. Estructura y conformaciones adoptadas por CaM.	7
Figura 1.4. Motivos de unión a CaM presentes en los diferentes CBDs.	9
Figura 1.6. Organización general de la familia de las IPKs.	11
Figura 1.7. Reacción catalizada por la IP <sub>3</sub> 3-K.	11
Figura 1.8. Organización general de la familia de las IP <sub>3</sub> 3-Ks.	13
Figura 1.9. Estructura de IP <sub>3</sub> 3-K.	15
Figura 1.10. Reacción catalizada por la IPMK.	16
Figura 1.11. Estructura de IPMK.	16
Figura 1.12. Reacciones catalizadas por la IP <sub>6</sub> -K.	17
Figura 1.13. Organización general de las IP <sub>6</sub> -Ks.	18
Figura 1.14. Reacciones catalizadas por la IP <sub>3</sub> 5/6-K.	19
Figura 1.15. Estructura de IP <sub>3</sub> 5/6-Ks.	20
Figura 1.16. Reacción catalizada por la IP <sub>5</sub> 2-K.	21
Figura 3.1. Etapas básicas de la cristalografía macromolecular de rayos X.	29
Figura 4.1. Purificación de IP <sub>5</sub> 2-K.	53
Figura 4.2. Pasos finales de la purificación de IP <sub>5</sub> 2-K.	54
Figura 4.3. Cristalización de IP <sub>5</sub> 2-K.	55
Figura 4.4. Patrones de difracción de IP <sub>5</sub> 2-K.	56
Figura 4.5. Representación de la función de autorrotación de los cristales de IP <sub>5</sub> 2-K/IP <sub>6</sub> /ADP (2 moléculas/u.a.).	58
Figura 4.6. Mapas de densidad electrónica de los cristales correspondientes a los complejos de IP <sub>5</sub> 2-K.	60
Figura 4.7. Organización general de la IP <sub>5</sub> 2-K de <i>A. thaliana</i> .	61
Figura 4.8. Organización del lóbulo N-terminal (N <sub>t</sub> ) de IP <sub>5</sub> 2-K.	62
Figura 4.9. Organización del lóbulo C-terminal (C <sub>t</sub> ) de IP <sub>5</sub> 2-K.	63
Figura 4.10. Interacciones del nucleótido en el centro activo de IP <sub>5</sub> 2-K.	64
Figura 4.11. Diagrama de interacciones entre IP <sub>5</sub> 2-K y sustratos.	65
Figura 4.12. Interacciones del inositido en el centro activo de IP <sub>5</sub> 2-K.	66
Figura 4.13. Interacción del bucle L3 con elementos de reconocimiento del IP.	67
Figura 4.14. Superficie electrostática de IP <sub>5</sub> 2-K.	68
Figura 4.15. Interacciones de los átomos de Mg <sup>2+</sup> en el centro activo.	69
Figura 4.16. Cambio de conformación del D407 debido a la presencia de AMPPNP.	70
Figura 4.17. Sitio de unión a Zn <sup>2+</sup> de IP <sub>5</sub> 2-K.	71
Figura 4.18. Purificación de mutantes GST-IP <sub>5</sub> 2-K en el sitio de Zn <sup>2+</sup> .	72
Figura 4.19. Alineamiento de IP <sub>5</sub> 2-Ks de diferentes especies.	73
Figura 4.20. Disposición de los bucles que albergan los diferentes motivos conservados implicados en la actividad enzimática.	74
Figura 4.21. Organización de la familia IPK.	77
Figura 4.22. Lóbulos de unión a IP de cada uno de los miembros de la familia IPK.	78
Figura 4.23. IP <sub>5</sub> 2-K vs IP <sub>3</sub> 3-K A.	78
Figura 4.24. IP <sub>5</sub> 2-K y proteína kinasa A.	80
Figura 4.25. Los elementos típicos de la conformación activa de PKs están ausentes en la IP <sub>5</sub> 2-K.	81
Figura 4.26. Estructura de IP <sub>5</sub> 2-K y cierre del centro activo.	84
Figura 4.27. SDS-PAGE de la muestra final de IP <sub>5</sub> 2-K <sub>W129A</sub> .	85
Figura 4.28. Cristalización de IP <sub>5</sub> 2-K <sub>W129A</sub> e IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP.	86
Figura 4.29. Patrones de difracción de IP <sub>5</sub> 2-K <sub>W129A</sub> .	87
Figura 4.30. Sección a 1σ del mapa de densidad electrónica de IP <sub>5</sub> 2-K <sub>W129A</sub> .	88

Figura 4.31. Comparación de las estructuras de la IP <sub>5</sub> 2-K en presencia y ausencia de inositido.	90
Figura 4.32. Superposición de las estructuras de IP <sub>5</sub> 2-K en ausencia/presencia de inositido.	91
Figura 4.33. El sitio de unión a inositido se encuentra preformado en la forma apo.	92
Figura 4.34. Superposición de las tres formas de IP <sub>5</sub> 2-K.	93
Figura 4.35. El sitio de unión al nucleótido en ausencia de inositido.	94
Figura 4.36. Cambios en los elementos $\alpha 6/L3$ .	95
Figura 4.37. Análisis de la secuencia de diferentes especies de IP <sub>5</sub> -2K ( <i>sequence logos</i> ).	97
Figura 4.38. Ensayos cinéticos de IP <sub>5</sub> 2-K.	98
Figura 4.39. Enantiómeros I(1,4,5,6)P <sub>4</sub> e I(3,4,5,6)P <sub>4</sub> .	99
Figura 4.40. Complejo IP <sub>5</sub> 2-K/I(3,4,5,6)P <sub>4</sub> /AMPPNP.	99
Figura 4.41. Ensayos de fluorescencia intrínseca de IP <sub>5</sub> 2-K.	101
Figura 4.42. Ensayos de unión a ligando de la IP <sub>5</sub> 2-K.	102
Figura 4.43. Representación fraccional de los cambios en la emisión de fluorescencia para la proteína nativa.	103
Figura 4.44. Gel nativo de las diferentes muestras de IP <sub>5</sub> 2-K.	103
Figura 4.45. SDS-PAGE de los ensayos de proteólisis limitada de IP <sub>5</sub> 2-K e IP <sub>5</sub> 2-K <sub>W129A</sub> .	104
Figura 4.46. Posible ciclo catalítico de la IP <sub>5</sub> 2-K.	106
Figura 4.47. SDS-PAGE mostrando la muestra pura del complejo IP <sub>3</sub> 3-K <sub>(158-183)</sub> /CaM.	108
Figura 4.48. Cristalización del complejo CaM/ IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	109
Figura 4.49. Mapas de densidad electrónica obtenidos de los diferentes cristales analizados.	113
Figura 4.50. Complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	114
Figura 4.51. Los residuos hidrofóbicos tienen un papel clave en la formación del complejo CaM/ IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	115
Figura 4.52. Bolsillo hidrofóbico 1 del complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	116
Figura 4.53. Bolsillo hidrofóbico 2 del complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	117
Figura 4.54. Bolsillo hidrofóbico 3 del complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	117
Figura 4.55. Interacciones polares del complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> .	118
Figura 4.56. Motivos de unión a CaM.	119
Figura 4.57. Comparación del CBD de IP <sub>3</sub> 3-K con un motivo de tipo 1-10.	121
Figura 4.58. Comparación del motivo de unión a CaM de IP <sub>3</sub> 3-K y un motivo de tipo 1-14.	123
Figura 4.59. Comparación del motivo de unión a CaM de IP <sub>3</sub> 3-K y un motivo de tipo 1-16.	125
Figura 4.60. El motivo de unión a CaM de la IP <sub>3</sub> 3-K A se encuadra dentro de una nueva clase de motivos 1-13-16.	126
Figura 4.61. Comparación entre las dos formas de CaM observadas en este trabajo.	128
Figura 4.62. Alineamiento entre las diferentes isoformas de la IP <sub>3</sub> 3-K de diferentes especies.	129
Figura 4.63. Perfil de dispersión promedio obtenido a partir de muestras del complejo CaM/IP <sub>3</sub> 3-K <sub>(128-461)</sub> .	130
Figura 4.64. Envoltura obtenida a partir de los datos de dispersión a bajo ángulo (SAXS) del complejo CaM/IP <sub>3</sub> 3-K <sub>(128-461)</sub> .	131
Figura 4.65. Modelo a alta resolución del complejo CaM/IP <sub>3</sub> 3-K <sub>(128-461)</sub> .	132
Figura 4.66. Modelo de la activación de IP <sub>3</sub> 3-K por CaM.	133
Figura 7.1. El cristal. a) Ejes y ángulos de una celdilla unidad.	162
Figura 7.2. Diagrama de fases del proceso de cristalización.	164
Figura 7.3. Variantes del método de cristalización por difusión de vapor.	165
Figura 7.4. Cristalización de alto rendimiento.	165
Figura 7.5. Ley de Bragg.	167
Figura 7.6. Generación y uso de radiación de sincrotrón.	168
Figura 7.7. Elementos para la realización de un experimento de difracción.	169

Figura 7.8. Patrón de difracción. Relación entre la resolución y la calidad de los mapas de densidad electrónica.	171
Figura 7.9. El problema de la fase.	172
Figura 7.10. Espectro de fluorescencia alrededor del pico de absorción del selenio.	174
Figura 7.11. Método del Reemplazamiento Molecular.	175

# Índice de tablas

Tabla 3.1 Condiciones empleadas para la derivatización con átomos pesados de IP <sub>5</sub> 2-K.	33
Tabla 3.2 Parámetros de la recogida de datos de los diferentes complejos de IP <sub>5</sub> 2-K.	33
Tabla 3.3 Parámetros de recogida de datos de IP <sub>5</sub> 2-K <sub>W129A</sub> y los complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP y IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP.	37
Tabla 3.4 Parámetros de recogida de datos de CaM/IP <sub>3</sub> 3-K CBD.	45
Tabla 3.5. Parámetros de recogida y procesamiento de datos de SAXS.	49
Tabla 4.1. Condiciones de cristalización de los diferentes complejos de IP <sub>5</sub> 2-K.	56
Tabla 4.2. Procesado de datos de los diferentes complejos de IP <sub>5</sub> 2-K.	57
Tabla 4.3. Procesado de datos del derivado con plomo de IP <sub>5</sub> 2-K.	58
Tabla 4.4. Estadísticas del refinamiento de los diferentes conjuntos de datos de IP <sub>5</sub> 2-K.	59
Tabla 4.5. Tasa de conversión de IP <sub>5</sub> .	71
Tabla 4.6. Actividad de los diferentes mutantes de IP <sub>5</sub> 2-K.	81
Tabla 4.7. Procesado de datos de cristales de IP <sub>5</sub> 2-K <sub>W129A</sub> y de complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP.	87
Tabla 4.8. Reemplazamiento Molecular de IP <sub>5</sub> 2-K <sub>W129A</sub> , IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP.	88
Tabla 4.9. Estadísticas del refinamiento de IP <sub>5</sub> 2-K <sub>W129A</sub> y los complejos IP <sub>5</sub> 2-K <sub>W129A</sub> /AMPNP e IP <sub>5</sub> 2-K <sub>W129A</sub> /ADP.	89
Tabla 4.10. Actividad enzimática relativa frente a IP <sub>5</sub> de IP <sub>5</sub> 2-K y diferentes mutantes.	96
Tabla 4.11. Parámetros cinéticos de IP <sub>5</sub> 2-K e IP <sub>5</sub> 2-K <sub>W129A</sub> .	98
Tabla 4.12. Discriminación de enantiómeros en IP <sub>5</sub> 2-K nativa y diferentes mutantes.	100
Tabla 4.13 Procesado de datos de CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> y CaM.	110
Tabla 4.14. Reemplazamiento Molecular de CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> y CaM.	111
Tabla 4.15. Estadísticas de refinamiento del complejo CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> y CaM.	112
Tabla 4.16. Parámetros estructurales obtenidos mediante SAXS.	130

# Índice de ecuaciones

Ecuación 7.1. Ley de Bragg.	167
Ecuación 7.2. Función de densidad electrónica.	171
Ecuación 7.3. Factor de dispersión atómico.	173
Ecuación 7.4. $R_{\text{work}}/R_{\text{free}}$ .	176

# Abreviaturas

ADP: adenosín difosfato

AMPPNP: adenosín 5'-( $\beta,\gamma$ -imido)trifosfato

ATP: adenosín trifosfato

Bis-tris: bis (2-hidroxietil)-amino-tris(hidroximetil)-metano

Bis-tris propano: 1,3-bis(tris(hidroximetil)metilamino)propano

CaM: calmodulina, derivado del acrónimo inglés ***Calcium Modulated Protein***

CaMKI: calmodulina Kinasa I

CaMKII: calmodulina Kinasa II

CaMKK: calmodulina Kinasa-Kinasa

CBD: dominio de unión a calmodulina, derivado del acrónimo inglés ***Calmodulin Binding Domain***

CCD: dispositivo de carga acoplada (del inglés **Charge Coupled Device**)

CMOS: semiconductor complementario de óxido de metal (del inglés **Complementary Metal Oxide Semiconductor**)

DAG: diacilglicerol

EDTA: ácido etilendiamino tetra-acético (del inglés **Ethylenediaminetetraacetic acid**)

EGTA: ácido etilenglicol tetra-acético (del inglés **Ethylene Glycol Tetraacetic Acid**)

DTT: **d**itiotreitol

ESRF: instalación europea de radiación sincrotrón (del inglés **European Synchrotron Radiation Facility**)

GST: glutatión-**S**-transferasa

HEPES: ácido 4-(2-hidroxietil)-1-piperazinaetanosulfónico (del inglés 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

I(1,4)P<sub>2</sub>: inositol 1, 4-bifosfato

I(1,4,5)P<sub>3</sub>: inositol 1, 4, 5-trifosfato

I(1,3,4)P<sub>3</sub>: inositol 1, 3, 4-trifosfato

I(1,3,4, 5)P<sub>4</sub>: inositol 1, 3, 4, 5-tetrakisfosfato

I(3,4,5,6)P<sub>4</sub>: inositol 3, 4, 5, 6-tetrakisfosfato

I(1,3,4,5,6)P<sub>5</sub>: inositol 1, 3, 4, 5, 6 pentakisfosfato

I(1,2,3,4,5,6)P<sub>6</sub> o IP<sub>6</sub>: inositol hexakisfosfato

InsP K: inositol fosfato kinasa

InsP Ps: inositol fosfato fosfatasas

INPP5A: inositol polifosfato 5 fosfatasa

IPTG: **I**sopropil-  $\beta$  -D-**t**iogalactósido

IP: inositol fosfato

IPK: inositol polifosfato kinasa

IPMK: inositol polifosfato multikinasa

IP<sub>3</sub> 3-K: inositol 1,4,5-trifosfato 3-kinasa

IP<sub>3</sub> 5/6-K: inositol 1,3,4-trisfosfato 5/6-kinasa

IP<sub>5</sub> 2-K: inositol 1,3,4,5,6-pentaquisfosfato 2-kinasa

IP<sub>6</sub>-K: inositol 1,2,3,4,5,6-hexaquisfosfato-kinasa

Lóbulo N<sub>t</sub>: lóbulo N-terminal

Lóbulo C<sub>t</sub>: lóbulo C-terminal

Lpa: del inglés *low phitic acid*

LSL: lectina de *Laetiporus sulphureus*. (del inglés *Laetiporus sulphureus* Lectin)

MAD: dispersión anómala múltiple (del inglés **M**ulti-wavelength **A**nomalous **D**ispersion)

MIR: remplazamiento isomorfo múltiple (del inglés **M**ultiple **I**somorphous **R**eplacement)



MR: Reemplazamiento Molecular (del inglés **M**olecular **R**eplacement)

mTOR: del inglés **m**ammalian **T**arget **O**f **R**apamicyn

SDS-PAGE: Electroforesis en gel de poliacrilamida con dodecilsulfato sódico

(del inglés **P**oly**a**crylamide **G**el **E**lectrophoresis- **S**odium **D**odecyl **S**ulfate)

PDB: banco de datos de estructuras de proteínas (del inglés **P**rotein **D**ata **B**ank)

PEG: polietilenglicol

PKA: proteína Kinasa A

PKC: proteína Kinasa C

PtdIns: fosfatidil-inositol

PtdIns-4-P: fosfatidil-inositol-4-fosfato

PtdIns(4, 5)P<sub>2</sub>: fosfatidil-inositol-4, 5-bifosfato

rmsd: desviación cuadrática media (del inglés **r**oot **m**ean **s**quare **d**eviation)

SAD: dispersión anómala simple (del inglés **S**ingle wavelength **A**nomalous

**D**ispersion)

SAXS: dispersión de rayos X a bajo ángulo (del inglés **S**mall **A**ngle **X**-ray **S**cattering).

*s*MLCK: kinasa de la cadena ligera de miosina del músculo liso (del inglés **s**mooth **M**yosin **L**ight **C**hain **K**inase)

Tris: tris-(hidroximetil) aminometano

u.a. :unidad asimétrica

5-PP-IP<sub>5</sub>: 5-pirofosfoinositol pentakisfosfato

# ABSTRACT



# Structural biology of inositol kinases, enzymes in charge of regulation of inositol phosphates.

## Introduction

Inositol is the building block for the synthesis of a large number of molecules in the cell. During the last years, the interest in these compounds has increased, enlarging the number in cloning and characterization of different kinases, phosphatases, lipases and effectors that regulate the levels and actions of these compounds. A total of 84 genes whose products are related to inositol metabolism have been characterized (York, 2006).

There are two different families of compounds containing the inositol ring, referred as inositides: inositol lipids and inositol phosphates.

Despite its low abundance, inositol lipids play a fundamental role in various cellular functions such as cell proliferation, apoptosis, motility, cytoskeletal regulation or vesicular trafficking, being the most studied inositides to date (Toker, 2002). Among them, the most important inositol lipid is PtdIns (4,5) P<sub>2</sub>. PtdIns (4,5) P<sub>2</sub> can be hydrolyzed by phospholipase C, generating two important second messengers: diacylglycerol (DAG) and inositol 1,4,5-triphosphate or IP<sub>3</sub> (Berridge, 1987). IP<sub>3</sub> is a soluble inositide from the second family, the inositol phosphates, that has two major roles in the cell: to promote Ca<sup>2+</sup> release from intracellular stores and to act as a starting point for the synthesis of a number of other inositol phosphates (Irvine y Schell, 2001).

These inositol polyphosphates are a wide group of second messengers, involved in key cellular events (Irvine y Schell, 2001). Their levels are tightly regulated by the action of different kinases and phosphatases. This thesis is focused on the structural biology of inositol phosphate kinases (IPKs). The IPKs have a catalytic domain with kinase activity and, some of them, other domains with different functions, such as regulation etc.

These kinases have been classified into several families according to the folding they present at the kinase domain (González *et al.*, 2004):

> Family I: Inositol polyphosphate kinases (IPKs): formed by IP<sub>3</sub> 3-kinases (IP<sub>3</sub> 3-Ks), IP multikinases (IPMKs) and IP<sub>6</sub> Kinases (IP<sub>6</sub> Ks). These enzymes present a protein kinase fold.

> Family II: Inositol (1,3,4) trisphosphate 5/6-Kinase with an ATP/GRASP fold.

> Family III: Inositol (1,3,4,5,6) pentakisphosphate 2 kinases (IP<sub>5</sub> 2-Ks), which have an unknown fold.

This work is focused on Inositol (1,3,4,5,6) pentakisphosphate 2 kinases and IP<sub>3</sub> 3-kinases.

IP<sub>5</sub> 2-K converts inositol pentakisphosphate (IP<sub>5</sub>) to inositol hexakisphosphate or phytic acid (IP<sub>6</sub>) using ATP as phosphate donor. IP<sub>6</sub> is present in all eukaryote cells and plays an essential role in processes such as lymphocyte development or apoptosis (Shi *et al.*, 2006). In addition, IP<sub>6</sub> tends to accumulate in plant seeds, what may have detrimental effects in human health and environment (Raboy, 2001). For example, phytic acid is a potent chelator, contributing to malnutrition in populations where the diet is grain-based. In spite of the broad interest of IP<sub>6</sub> in biomedicine and biotechnology, there is a lack in structural information concerning this enzyme.

On the other hand, inositol 1,4,5-trisphosphate 3-kinase (IP<sub>3</sub> 3-K) catalyzes the phosphorylation of IP<sub>3</sub> yielding inositol 1,3,4,5-tetrakisphosphate (IP<sub>4</sub>) which is involved in T lymphocyte development. Soluble IP<sub>3</sub> binds directly to endoplasmic reticulum-resident calcium channels and regulates their permeability establishing an intracellular flux of calcium in the cell (Pattani y Banting, 2004). Calmodulin (Yañez *et al.*, 2012), in presence of calcium, binds IP<sub>3</sub> 3-K increasing its activity and, in turn, reducing intracellular IP<sub>3</sub> concentration and terminating IP<sub>3</sub> action. This IP<sub>3</sub> role has been studied for years, being considered as a paradigm of second messenger action in cells (Berridge, 2009). The structure of IP<sub>3</sub> 3-K kinase domain was reported in 2004 (González *et al.*, 2004; Miller y Hurley, 2004). However there is a lack of knowledge about the enzymatic regulation of IP<sub>3</sub> 3-K by calmodulin, with the exception of some assays to identify the IP<sub>3</sub> 3-K region involved in CaM binding (Takazawa y Erneux, 1991; Erneux *et al.*, 1993). To date, no structural information about the IP<sub>3</sub> 3-K/CaM complex is available.

## Objetives

During the last years, there has been a breakthrough in the characterization of InsP Ks because of their importance in different physiological processes. Despite of this, there is still a lack of knowledge about important aspects of these enzymes. For example, no structural information about IP<sub>5</sub> 2-Ks (family III) is available. On the other

hand, the structural determinants of the regulation of IP<sub>3</sub> 3-K (family I) by CaM remain elusive.

Therefore, the objectives of this work have been:

- IP<sub>5</sub> 2-K structure determination and analysis of its structure-function relationship:

1. Cloning, expression, purification, crystallization and structural determination of IP<sub>5</sub> 2-K from *A. thaliana*.

2. Structural and functional analysis of the IP<sub>5</sub> 2-K from *A. thaliana*. Elucidation of the structural determinants that govern the binding of substrates (IP<sub>5</sub>, ATP) and products (IP<sub>6</sub>, ADP).

3. Characterization of the structural changes during the catalytic cycle of IP<sub>5</sub> 2-K.

- Structure determination of the complex between CaM binding domain of human IP<sub>3</sub> 3-KA (CBD IP<sub>3</sub> 3-K) and human CaM:

1. Purification of the Calmodulin binding domain of IP<sub>3</sub> 3-K and CaM complex.

2. Crystallization and structural determination of the complex.

3. Analysis of the structural determinants that govern the binding of CaM and IP<sub>3</sub> 3-K.

4. Building a model of the complex between CaM and IP<sub>3</sub> 3-K<sub>(128-461)</sub>, by combining data at high and low resolution.

## Metodology. Materials and Methods

X-ray crystallography is the most powerful technique for determining three dimensional structures of molecules at atomic resolution, independently of its size. The technique allows the study of peptides and proteins from small size to complex macromolecular machines.

In general, the structural determination of proteins by X-ray crystallography involves successive stages: cloning of the target gene, expression of the protein, purification, crystallization, data collection and structural determination. During the last years a number of techniques have been implemented to increase the performance of the different steps, including the use of robots for crystallization and the improvements associated to the development of 3<sup>rd</sup> generation synchrotrons. The development of easy-to-use programs has increased the number of protein structures determined during the

last years. Despite this automation several stages are considered still "bottlenecks". Thus, for example, obtaining a pure and homogenous sample that make posible the crystallization step, or the calculation of phases require extensive experimental work in order to successfully address the structural study (Rupp, 2009).

X-ray crystallography (including standard cloning techniques, protein expression and protein purification) has been used as main technique for the development of this thesis. In addition a number of other biochemical and biophysical assays has been used to fulfil the proposed objectives: small angle X-ray scattering (SAXS), site directed mutagenesis, fluorescence analysis, limited proteolysis and analysis employing native gels are among the complementary techniques employed in this work.

## Results

### IP<sub>5</sub> 2-K: complexes with substrates (IP<sub>5</sub> IP<sub>5</sub>+AMPPNP and IP<sub>4</sub>+AMPPNP) and products (IP<sub>6</sub> and IP<sub>6</sub>+AMPPNP).

IP<sub>5</sub> 2-K from *Arabidopsis thaliana* was cloned into a vector, expressed in *Escherichia coli* as fusion protein and purified. Crystals appeared always in the presence of substrate (IP<sub>5</sub> or IP<sub>4</sub>) or product (IP<sub>6</sub>). The best crystals obtained for X-ray crystallographic analysis belonged to space group P212121, with unit cell parameters a = 58.124, b = 113.591, c = 142.478 Å (2.2 Å resolution). A derivative with lead acetate was prepared employing IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP crystals, the structure of this complex being solved by MAD (3.2 Å resolution). Employing this model, the complexes of IP<sub>5</sub> 2-K with substrates (IP<sub>5</sub> 2-K/IP<sub>4</sub>/AMPPNP), (IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP) and products (IP<sub>5</sub> 2-K/IP<sub>6</sub>) were solved by Molecular Replacement and Fourier Difference Analysis.

IP<sub>5</sub> 2-K from *A. thaliana* folds into an  $\alpha + \beta$  structure composed of two lobes, N and C-terminal, connected by a loop (hinge). The ATP binding site, conserved from protein kinases, is located at the interface formed by the  $\beta$  sheets of the two lobes, residues from both lobes being involved in binding of the nucleotide. The inositide binding site is a region located at the  $\alpha$  helical part of the C-terminal lobe, (the C<sub>IP</sub> lobe). Most critical residues for binding and catalysis come from this region, although the N-terminal lobe also provides some key residues. The relevance of many of these residues was studied by site directed mutagenesis and activity assays, the results confirming the roles inferred by the structure. The protein has also a Zn<sup>2+</sup> binding site, which probably is involved in the correct folding of the enzyme.

This work has shown that, unexpectedly, the IP<sub>5</sub> 2-K belongs to the family of IPKs, although it is the most distant member of this family.

## IP<sub>5</sub> 2-K: apo form and complexes with nucleotide (AMPPNP and ADP)

Due to the failure of IP<sub>5</sub> 2-K to crystallize in absence of inositide, the putative conformational changes upon substrate binding were studied. Residues involved in a region forming a “clasp” over the active site (regions  $\alpha 6$  and L3) were mutated. One of the mutants (W129A) was expressed, purified, crystallized and solved, allowing the analysis of the IP<sub>5</sub> 2-K apo form and nucleotide bound form. The IP<sub>5</sub> 2-K apo form structure displays an open conformation, whereas the nucleotide-bound form shows a half-closed conformation, in contrast to the inositide-bound form obtained previously, which presents a closed conformation. Both, nucleotide and inositide binding, produce a rigid domain movements between N and C lobes, although local changes are also observed, especially in the “clasp” area. In addition, intrinsic fluorescence experiments were done. Changes in intrinsic fluorescence upon nucleotide and inositide binding are in agreement with the crystallographic data, as well as the limited proteolysis and native gel based assays.

In addition, a study of IP<sub>5</sub> 2-K discrimination between inositol 1,3,4,5-tetrakisphosphate and 3,4,5,6-tetrakisphosphate enantiomers was performed based on activity measurements of the different mutants and the previously obtained structure of the IP<sub>5</sub> 2-K/IP<sub>4</sub>/AMPPNP complex.

## IP<sub>3</sub> 3-K regulation by calmodulin

As mentioned before, although there is a detailed structural knowledge of the kinase domain of the three IP<sub>3</sub>-3K isoforms, no structural information is available about the interaction of IP<sub>3</sub> 3-K with Ca<sup>2+</sup>/calmodulin. During this work, the complex between human calmodulin and the calmodulin binding domain of human IP<sub>3</sub>3-K isoform A (residues 158 to 183; IP<sub>3</sub> 3-K<sub>(158-183)</sub>) was produced and crystallized in two different habits. Crystals from habit 1 belong to I432 space group with unit-cell dimensions a=b=c= 159.25 Å (2.34 Å resolution). These crystals contain one molecule of the complex. Crystals from habit 2 belong to space group P3221, with the following cell parameters a=b= 40.27 Å c=175.31 Å (1.84 Å resolution). Unfortunately, the latter were composed only of one molecule of calmodulin, the IP<sub>3</sub> 3-K calmodulin binding domain being absent.



The complex between CaM and IP<sub>3</sub> 3-K<sub>(158-183)</sub> shows a compact structure with ellipsoidal shape. The structure reveals that, as in other similar complexes, the peptide corresponding to IP<sub>3</sub> 3-K (158-183) is attached to both domains of CaM.

The CBD of IP<sub>3</sub> 3-K shows two different zones: a first region located in the N-term, which is forming a short 10 residue  $\alpha$  helix (residues 165-175), and a second zone at the C-term, which is forming a hairpin (residues 176-183), packed against the end of the  $\alpha$  helix (residues 182-183). CBD sequence contains a large number of residues with hydrophobic character: Trp167, Ile170, Met173, Val174, Leu176, Val178, Ile179 and Phe182

A typical CBD present two or three hydrophobic residues, separated by a variable distance. These residues are key in the interaction between the CBD and calmodulin via different calmodulin conserved pockets. Three different pockets are present in the CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> complex: pocket 1, in charge of Trp167 binding, pocket 2, in charge of Ile179 binding and pocket 3, involved in the binding of Phe182.

The relative position of these hydrophobic residues is an essential feature that characterizes the different CBD classes. To date, three different CBD classes have been described (1-10, 1-14 and 1-16). During this work, the structure of the complex obtained was compared with representatives of these three classes. This comparison has revealed that the IP<sub>3</sub> 3-K calmodulin-binding domain belongs to a novel 1-13-16 class.

To get insights into the IP<sub>3</sub> 3-K activation by calmodulin, a model of the complex between IP<sub>3</sub> 3-K<sub>(128-461)</sub> and calmodulin was obtained, combining data from small angle scattering (SAXS) and X-ray crystallography. This model allowed us to propose a model of activation of IP<sub>3</sub> 3-K by CaM.

## Conclusions

### IP<sub>5</sub> 2-K

1. Binary (with inositide) and ternary (with inositide and nucleotide) complexes of IP<sub>5</sub> 2-K from *A. thaliana* have been characterized. The enzyme has a  $\alpha + \beta$  fold composed of two lobes (N and C-terminal) connected by a hinge, being the IP binding lobe a novel feature.
2. The ATP binding site is located at the interface formed by the  $\beta$  sheets of the N and C -terminal lobes. Residues from both lobes are involved in binding the nucleotide.

3. Mainly a  $\alpha$  helix region located in the C-terminal lobe, called CIP lobe, forms the IP binding site. Most critical residues for binding IP<sub>5</sub> and IP<sub>6</sub> come from this region, although the N-terminal lobe also provides some key residues.
4. The molecular mechanisms of inositide and nucleotide recognition, as well as the enzymatic mechanism involved in the production of phytic acid have been characterized. The role of different residues has been confirmed by site directed mutagenesis experiments.
5. A Zn<sup>2+</sup> binding site has been characterized. This metal is probably involved in the correct folding of IP<sub>5</sub> 2-K.
6. IP<sub>5</sub> 2-K belongs to the family of IPKs, although it is the most distant member of the family. The folding of the N-terminal lobe and the ATP recognition pattern is conserved with the protein kinase family.
7. It has been determined the crystal structure of the free form of IP<sub>5</sub> 2-K of *A. thaliana*, which has a remarkably open conformation compared to complexes crystallized in presence of inositide.
8. The crystal structure of the nucleotide bound form of IP<sub>5</sub> 2-K from *A. thaliana* has been determined. This crystal form shows an intermediate conformation between open (free) and closed (bound to inositide).
9. We have determined the role of the N-terminal lobe in the binding of substrates, and the molecular basis explaining the discrimination of enantiomeric substrates.
10. A mechanism explaining the catalytic cycle of IP<sub>5</sub> 2-K has been proposed by means of various biochemical assays in combination with the structural information available.

## IP<sub>3</sub> 3-K

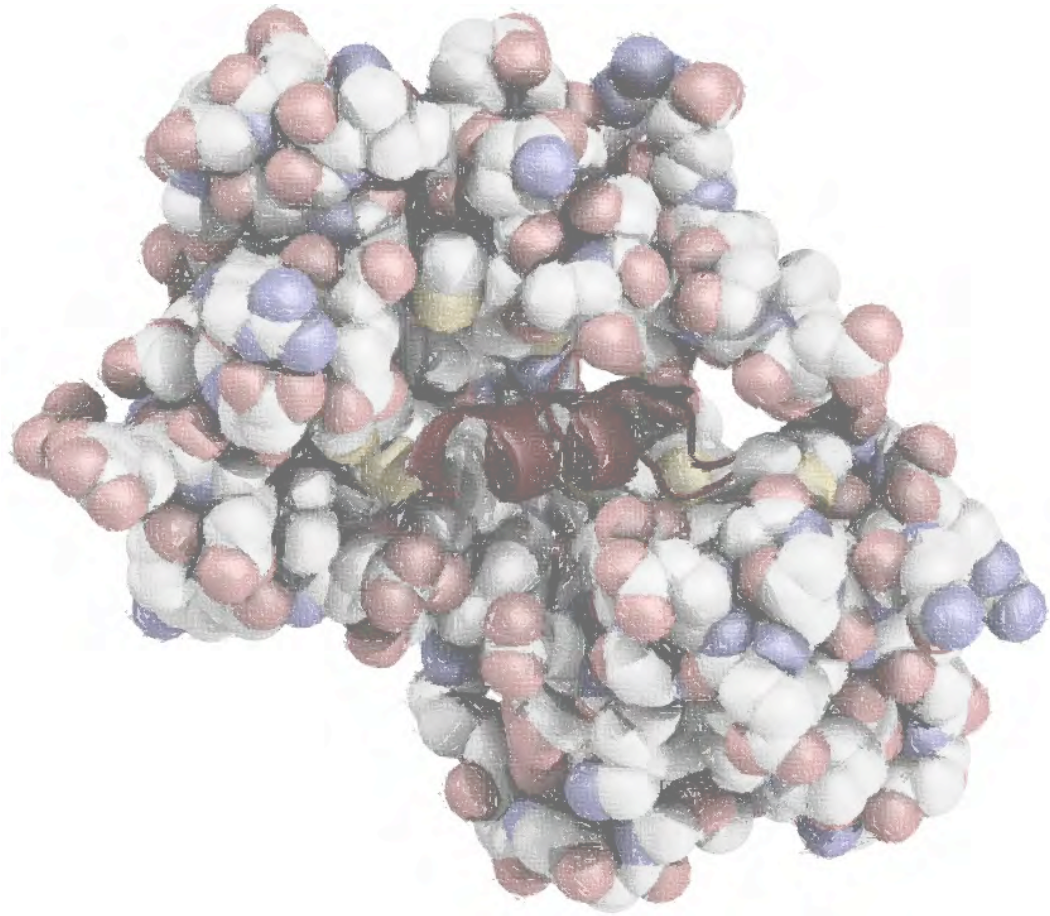
1. The complex between the Calmodulin binding domain of IP<sub>3</sub> 3-K A and CaM from *H. sapiens* has been purified, crystallized and the crystal structure has been elucidated by X-ray crystallography.
2. The structural determinants modulating the binding of these two proteins have been characterized.

3. A new class of CaM binding motif has been characterized (motif 1-13-16).
4. The combination of data from high and low resolution allowed us to propose a model for the activation of IP<sub>3</sub> 3-K by CaM.

## Bibliography

- Berridge M.J.** (1987). Inositol trisphosphate and diacylglycerol: two interacting second messengers. *Annu Rev Biochem.* **56** 159-193.
- Berridge M.J.** (2009). Inositol trisphosphate and calcium signalling mechanisms. *Biochim Biophys Acta.* **1793** (6): 933-940.
- Erneux C., Moreau, C., Vandermeers, A. y Takazawa, K.** (1993). Interaction of calmodulin with a putative calmodulin-binding domain of inositol 1,4,5-trisphosphate 3-kinase. Effects of synthetic peptides and site-directed mutagenesis of Trp165. *Eur J Biochem.* **214** (2): 497-501.
- González B., Schell, M.J., Letcher, A.J., Veprintsev, D.B., Irvine, R.F. y Williams, R.L.** (2004). Structure of a human inositol 1,4,5-trisphosphate 3-kinase: substrate binding reveals why it is not a phosphoinositide 3-kinase. *Mol Cell.* **15** (5): 689-701.
- Irvine R.F. y Schell, M.J.** (2001). Back in the water: the return of the inositol phosphates. *Nat Rev Mol Cell Biol.* **2** (5): 327-338.
- Miller G.J. y Hurley, J.H.** (2004). Crystal structure of the catalytic core of inositol 1,4,5-trisphosphate 3-kinase. *Mol Cell.* **15** (5): 703-711.
- Pattni K. y Banting, G.** (2004). Ins(1,4,5)P<sub>3</sub> metabolism and the family of IP<sub>3</sub>-3Kinases. *Cell Signal.* **16** (6): 643-654.
- Raboy V.** (2001). Seeds for a better future: 'low phytate' grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci.* **6** (10): 458-462.
- Rupp B.** (2009). Biomolecular Crystallography : Principles, Practice, and Application to Structural Biology. Garland Science.
- Shi Y., Azab, A.N., Thompson, M.N. y Greenberg, M.L.** (2006). Inositol phosphates and phosphoinositides in health and disease. *Subcell Biochem.* **39** 265-292.
- Takazawa K. y Erneux, C.** (1991). Identification of residues essential for catalysis and binding of calmodulin in rat brain inositol 1,4,5-trisphosphate 3-kinase. *Biochem J.* **280** ( Pt 1) 125-129.
- Toker A.** (2002). Phosphoinositides and signal transduction. *Cell Mol Life Sci.* **59** (5): 761-779.
- Yañez M., Gil-Longo, J. y Campos-Toimil, M.** (2012). Calcium binding proteins. *Adv Exp Med Biol.* **740** 461-482.
- York J.D.** (2006). Regulation of nuclear processes by inositol polyphosphates. *Biochim Biophys Acta.* **1761** (5-6): 552-559.





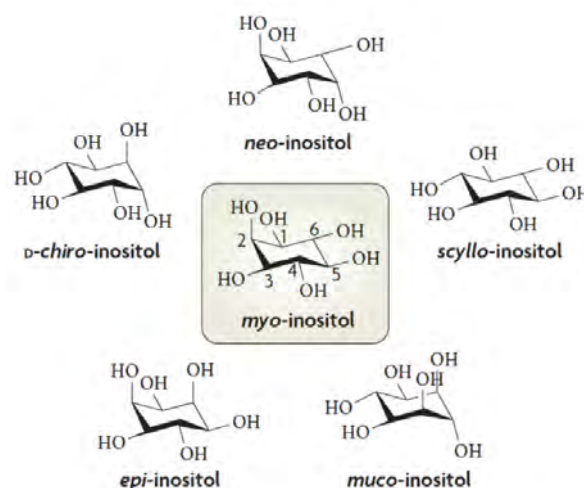
# 1 INTRODUCCIÓN





## 1.1 Los inositidos

El inositol es un ciclohexanohexol, es decir, un carbohidrato cíclico de seis carbonos con un total de seis grupos hidroxilo (uno por cada carbono) (**Figura 1.1**). El compuesto fue descrito por primera vez en el siglo XIX (Scherer, 1850). Los grupos hidroxilo del inositol pueden presentarse en dos posiciones posibles, axial o ecuatorial. La combinación de estas orientaciones para cada grupo hace posible la existencia de nueve posibles isómeros del inositol: *scyllo*, *myo*, *epi*, *chiro* (*D* o *L*), *neo*, *allo*, *cis* y *muco*. Seis de estos isómeros están presentes en la naturaleza (*scyllo*, *myo*, *D-chiro*, *L-chiro*, *neo* y *muco*) (**Figura 1.1**), siendo el *myo*-inositol el más abundante (Shi *et al.*, 2006). El *myo*-inositol presenta 5 de sus grupos hidroxilo en posición ecuatorial, encontrándose únicamente el hidroxilo-2 en posición axial.



**Figura 1.1. Isómeros del inositol.** Se representan los seis isómeros del inositol presentes en los seres vivos.

El inositol es la base para la síntesis de un gran número de moléculas en la célula, muchas de ellas con funciones clave. Según se han ido elucidando las funciones de estos compuestos, el interés en ellos ha ido en aumento, sucediéndose la clonación y caracterización de diferentes kinasas, fosfatasa, lipasa y efectores que regulan los niveles y acciones de estos metabolitos. De hecho se han caracterizado hasta un total de 84 genes cuyos productos están relacionados con el metabolismo de los inositoles (York, 2006).



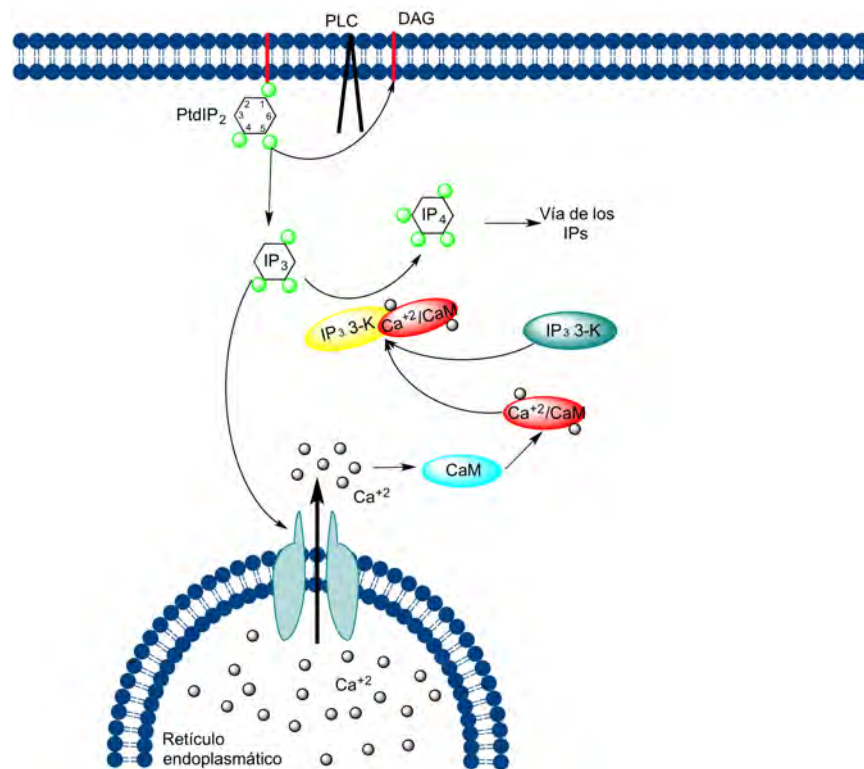
Dos son las familias más destacadas de compuestos basadas en el anillo de inositol: los inositoles lipídicos y los inositol fosfatos.

### 1.2 Inositoles lipídicos

Los inositoles lipídicos son una familia de moléculas que, pese a su escasa abundancia, juegan un rol fundamental en diversas funciones celulares como proliferación celular, apoptosis, motilidad, regulación del citoesqueleto o tráfico vesicular (Toker, 2002).

El fosfatidilinositol (PtdIns) es el inositol lipídico más abundante en las células de mamíferos (10% del total de los lípidos presentes en la membrana plasmática). El compuesto contiene en una cabeza polar de *myo*-inositol, la cual está unida mediante un enlace éster al fosfolípido correspondiente. La cabeza polar de inositol de este compuesto presenta 5 posiciones susceptibles de ser fosforiladas, aunque *in vivo* solo se han observado fosforiladas tres de ellas (las posiciones 3, 4 y 5), lo cual permite la existencia de siete inositoles lipídicos diferentes en la membrana celular (Toker, 2002). La fosforilación y desfosforilación de estas posiciones se produce mediante diferentes kinasas y fosfatasas, siendo el PtdIns el sustrato para la síntesis del resto de inositoles lipídicos (denominados fosfoinosítidos). La vía de síntesis de fosfoinosítidos más común comienza con la fosforilación de este compuesto en la posición 4 mediante la PI 4-kinasa para generar PtdIns-4-P. A continuación, la actuación de una PtdIns-4-P 5 kinasa sintetiza el PtdIns(4,5)P<sub>2</sub> (Payraastre *et al.*, 2001). Este compuesto es de capital importancia, pese a ser muy escaso en la membrana (1% sobre su composición total) pues por sí solo es capaz de regular diferentes proteínas de la superficie celular, así como de reclutarlas hacia la membrana plasmática mediante los dominios de tipo plectrina (PH), FYVE, o PX que estas poseen (Lemmon, 2003). Muchas de estas proteínas están relacionadas con tráfico de vesículas (Liscovitch y Cantley, 1995) o remodelación del citoesqueleto de actina (Janmey, 1994).

Además de estas funciones, el PtdIns(4,5)P<sub>2</sub> puede ser hidrolizado por la fosfolipasa C, o bien ser fosforilado de nuevo. En la hidrólisis mediante la fosfolipasa C, se generan dos importantes segundos mensajeros: el diacilglicerol (DAG) y el inositol 1,4,5-trifosfato o IP<sub>3</sub> (Berridge, 1987). El DAG queda anclado a la membrana plasmática, regulando diferentes enzimas (como por ejemplo la proteína kinasa C (Zeng *et al.*, 2012). El IP<sub>3</sub>, compuesto soluble, es liberado al citosol, donde esta sencilla molécula adquiere dos papeles fundamentales: promover la liberación de Ca<sup>2+</sup> desde las reservas intracelulares (Berridge y Irvine, 1984) por un lado, y actuar como punto de inicio para la síntesis de otra serie de mensajeros intracelulares: los inositol fosfatos (Irvine y Schell, 2001) (**Figura 1.2**).



**Figura 1.2. Funciones del  $\text{I}(1,4,5)\text{P}_3$  en la célula.** La hidrólisis del  $\text{PtdIns}(4,5)\text{P}_2$  provoca la liberación de  $\text{IP}_3$  y  $\text{DAG}$ . El aumento de  $\text{IP}_3$  provoca la liberación de  $\text{Ca}^{2+}$  del retículo endoplasmático. El  $\text{Ca}^{2+}$  activa a  $\text{CaM}$ , que a su vez se une a  $\text{IP}_3$  3-K activándola.  $\text{IP}_3$  3-K cataliza la síntesis de  $\text{I}(1,3,4,5)\text{P}_4$  (punto de inicio para la síntesis del resto de IPs). Al disminuir el  $\text{I}(1,4,5)\text{P}_3$  finaliza la liberación de  $\text{Ca}^{2+}$  desde el retículo.

## 1.3 Liberación de $\text{IP}_3$ y señalización por $\text{Ca}^{2+}$

### 1.3.1 Introducción

El  $\text{IP}_3$  es uno de los mensajeros intracelulares más conocidos y mejor estudiados, siendo un ejemplo paradigmático de segundo mensajero. La principal función del  $\text{Ins}(1,4,5)\text{P}_3$  es la movilización de  $\text{Ca}^{2+}$  (**Figura 1.2**) (Berridge y Irvine, 1984). Para ejercer este efecto, este inositol fosfato se une a receptores localizados en la membrana del retículo endoplasmático, que son canales iónicos que permiten el paso de  $\text{Ca}^{2+}$  a su través (Mikoshiba, 2007) (**Figura 1.2**).

Un incremento en los niveles de  $\text{Ca}^{2+}$  libre en el citosol representa la señal de respuesta a multitud de estímulos extracelulares. Estos estímulos y sus correspondientes respuestas, son muy diferentes según el tipo celular. Por ejemplo, el desarrollo embrionario se inicia mediante un aumento repentino de la concentración de calcio

intracelular, el cual está desencadenado por la entrada del espermatozoide en el óvulo. Otros ejemplos típicos son el inicio de la contracción en células musculares, o la inducción de la secreción de diferentes compuestos en células secretoras (Berridge, 2009).

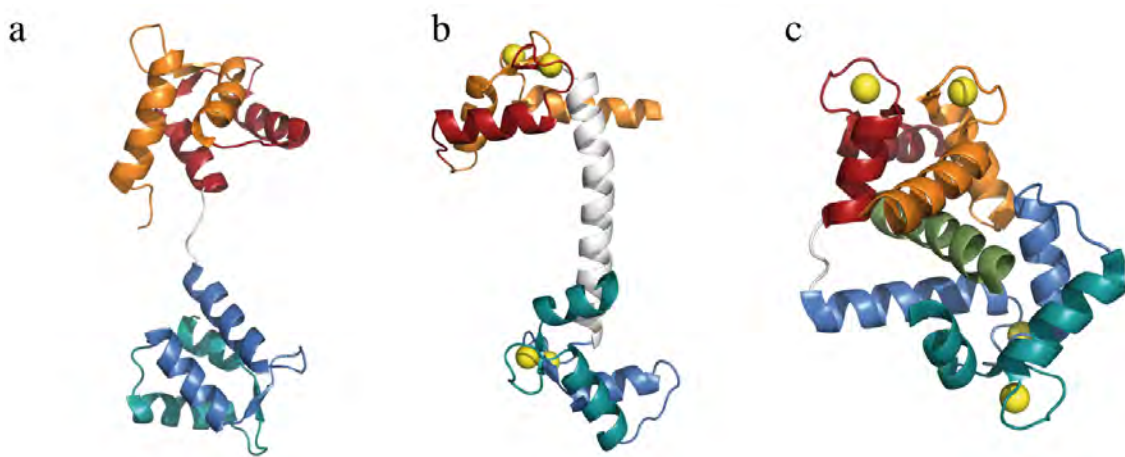
### 1.3.2 El $\text{Ca}^{2+}$ como segundo mensajero. Proteínas de unión a $\text{Ca}^{2+}$

¿Cómo los cambios en la concentración de calcio intracelular pueden producir cambios intracelulares? Por un lado, existen diferentes proteínas que son reguladas mediante unión directa del  $\text{Ca}^{2+}$  (por ejemplo la proteína kinasa C) (Zeng *et al.*, 2012), mientras que otras actúan como intermediarios. Entre estos intermediarios existe una amplia familia de proteínas que se distinguen por contar con el denominado “motivo EF” o “mano EF” (Kretsinger y Nockolds, 1973). El motivo consiste en una hélice N-terminal (hélice E) y una hélice C-terminal (hélice F), dispuestas de forma perpendicular y unidas mediante un bucle encargado de la unión del calcio (Yañez *et al.*, 2012). En general, las proteínas que actúan como sensor de  $\text{Ca}^{2+}$  pueden responder de dos formas: bien uniendo  $\text{Ca}^{2+}$  sin que este induzca en ellas un cambio conformacional significativo de forma que estas proteínas actúan como “tampón” o transportadores de  $\text{Ca}^{2+}$ , o bien induciendo un cambio conformacional, actuando entonces la proteína como “sensor” de  $\text{Ca}^{2+}$ . Como ejemplo típico de este último grupo está la calmodulina (Chin y Means, 2000).

### 1.3.3 La calmodulina, un sensor de calcio prototipo

La calmodulina (del acrónimo del inglés ***Calcium Modulated Protein; CaM***) es una de las proteínas de unión a  $\text{Ca}^{2+}$  mejor caracterizadas, estando presente en todos los organismos eucariotas (Means *et al.*, 1982). Consta de 148 residuos y se compone de dos pares de manos EF (un par en el N-terminal, manos 1 y 2, y un par en el C-terminal, manos 3 y 4) unidos por un conector flexible (**Figura 1.3a**). Los dos dominios N y C-terminal poseen una elevada homología en secuencia (75%), así como gran similitud a nivel estructural. La proteína puede adoptar diferentes conformaciones, según los elementos a los que esté unida. Cuando la proteína no está unida a  $\text{Ca}^{2+}$  (apocalmodulina) los dominios N y C-terminal presentan una conformación “cerrada” o “semicerrada”, con las hélices empaquetadas y ambos dominios conectados por un bucle (Kuboniwa *et al.*, 1995). En la segunda de las conformaciones (complejo  $\text{Ca}^{2+}/\text{CaM}$ ) ambos dominios pasan a una forma abierta, estando conectados ambos mediante una larga hélice  $\alpha$ . Además, una gran cantidad de residuos hidrofóbicos (la mayoría metioninas) quedan expuestos (**Figura 1.3b**) (Taylor *et al.*, 1991). Este cambio

se produce a concentraciones de calcio del orden de submicromolar, lo cual hace que la proteína pueda actuar como sensor de calcio. Esta conformación de la calmodulina posee una gran flexibilidad, lo que le permite adoptar una tercera conformación. Esta se da cuando la CaM se une a las proteínas que regula, consistiendo en el cierre de ambos dominios de la CaM alrededor de un segmento de la proteína objeto de regulación (Ikura *et al.*, 1992; Meador *et al.*, 1992) (**Figura 1.3c**). Estos segmentos son los denominados dominios de unión a CaM o CBDs (del inglés *Calmodulin Binding Domain*), de los que se hablará más adelante (Chin y Means, 2000). La CaM regula la función de más de un centenar de proteínas, las cuales participan en división celular, diferenciación, transcripción o síntesis de DNA, entre otras (Yañez *et al.*, 2012).

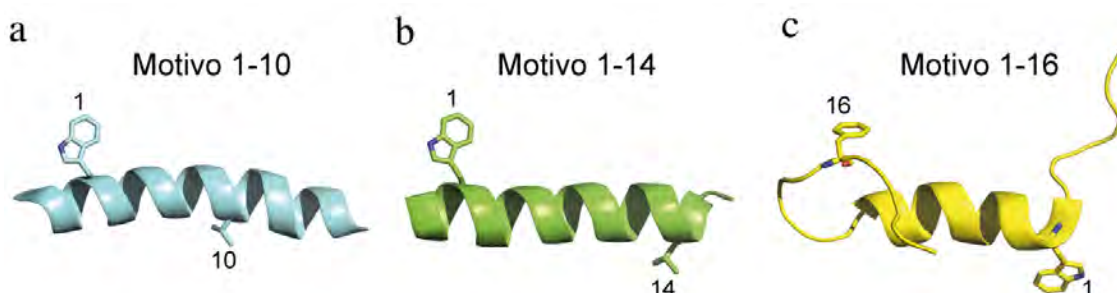


**Figura 1.3. Estructura y conformaciones adoptadas por CaM.** El dominio N-terminal se muestra con colores naranjas y rojos, mientras que el dominio C-terminal se muestra en tonos azules. El conector entre ambos dominios se muestra en blanco. Los iones  $\text{Ca}^{2+}$  se muestran mediante esferas amarillas. a) ApoCaM. b)  $\text{Ca}^{2+}$ /CaM. c) Complejo entre CaM y un CBD. El CBD se muestra en color verde.

#### 1.3.4 Proteínas reguladas por CaM. Dominios de unión a CaM (CBDs)

Existen diversos factores que posibilitan la regulación por CaM de un número tan elevado de proteínas. Por un lado los CBDs de las diferentes proteínas presentan características comunes: son fragmentos de aproximadamente 20-30 residuos de longitud, con un elevado número de residuos hidrofóbicos, que permiten una fácil unión a CaM cuando ésta tiene sus residuos hidrofóbicos expuestos (O'Neil y DeGrado, 1990). Por otro lado la presencia del conector que une los dominios N-terminal y C-terminal de la CaM dota a esta de una gran flexibilidad. Esta flexibilidad posibilita el cambio conformacional para unir los CBDs.

Los residuos hidrofóbicos son por tanto un elemento clave de los CBDs. Pese a tener un número variable de residuos hidrofóbicos en su secuencia, dos son los residuos que actúan como “anclaje principal” en la interacción, uno en el N-terminal de la secuencia del CBD, y otro en el C-terminal. Los dominios de unión a CaM se han clasificado en función de la distancia entre estos dos aminoácidos (Rhoads y Friedberg, 1997; Yap *et al.*, 2000). Por tanto, existen tres tipos fundamentales de dominios de unión a CaM: el dominio de unión de tipo 1-10 (Meador *et al.*, 1993) (**Figura 1.4a**), el dominio de unión de tipo 1-14 (Meador *et al.*, 1992) (**Figura 1.4b**) y el dominio de unión de tipo 1-16 (Osawa *et al.*, 1999) (**Figura 1.4c**).



**Figura 1.4. Motivos de unión a CaM presentes en los diferentes CBDs.** Los residuos clave de cada motivo se encuentran señalados. a) Motivo 1-10. b) Motivo 1-14. c) Motivo 1-16.

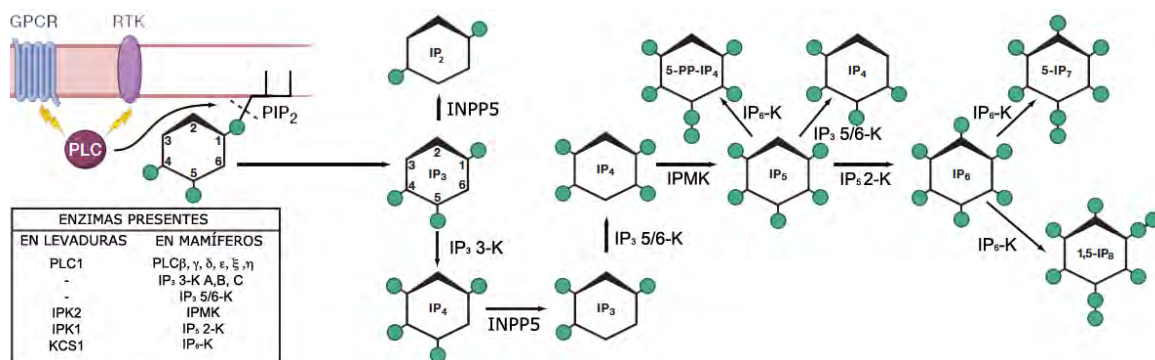
Una de las aproximaciones más frecuentemente empleadas a la hora de caracterizar estructuralmente la interacción entre diferentes proteínas y la CaM es el uso de péptidos cortos basados en la secuencia correspondiente del CBD de interés. Este tipo de aproximación, que ha sido validada experimentalmente (Kranz *et al.*, 2002), es la más habitual, existiendo muy pocos ejemplos en los que se haya podido caracterizar la interacción entre ambas proteínas completas (uno de estos ejemplos es la exotoxina del ántrax en complejo con CaM (Drum *et al.*, 2002)).

Una de las múltiples proteínas reguladas por unión a CaM es la inositol trifosfato 3 kinasa (IP<sub>3</sub> 3-K) (Takazawa *et al.*, 1990). Esta enzima es la encargada de metabolizar el IP<sub>3</sub>, finalizando por lo tanto la señal de liberación del Ca<sup>2+</sup>, y es, además, el punto inicial de una nueva vía de señalización formada por inositoles solubles: la vía de los inositol fosfatos, que se explicará en detalle a continuación.

## 1.4 Una nueva vía de señalización celular a partir de $IP_3$ : la vía de los inositol fosfatos

### 1.4.1 Introducción

Los inositol fosfatos (IPs) son un amplio grupo de moléculas (existen un total de 63 inositol fosfatos posibles (York, 2006)), cuya existencia se conoce desde hace casi un siglo (Posternak, 1919). Los diferentes inositol fosfatos están implicados de forma directa en una gran cantidad de funciones celulares claves: regulación de canales iónicos, transcripción, desarrollo embrionario, reparación de DNA o apoptosis entre otras (Irvine y Schell, 2001; Hatch y York, 2010; Saiardi, 2012; Shears *et al.*, 2012). Además pueden ejercer sus acciones de forma indirecta, bien actuando como cofactores de ciertas enzimas (Hatch y York, 2010), bien compitiendo con los inositoles lipídicos por los sitios de unión a estos de ciertas proteínas (Komander *et al.*, 2004).



**Figura 1.5. Vía de los inositol fosfatos.** En la ruta se representan solo las reacciones que se producen en mamíferos. Imagen modificada de (Hatch y York, 2010).

Como ya se ha indicado, la señalización mediante IPs comienza al romperse el  $PtdIns(4,5)P_2$  unido a la membrana mediante la acción de la fosfolipasa C, generándose  $IP_3$  y diacilglicerol (DAG). La adición sucesiva de grupos fosfatos al  $IP_3$  posibilita la aparición de una amplia red de IPs, cuyos niveles están regulados por diferentes Inositol Fosfato Kinasas (InsP Ks) e Inositol Fosfato Fosfatasa (InsP Ps) (**Figura 1.5**). La ruta de síntesis de los diferentes inositol fosfatos presenta variaciones en cada organismo. Por ejemplo, en el caso de mamíferos la acción sucesiva de cuatro quinasas y una fosfatasa es necesaria para la obtención de inositol hexakisfosfato ( $I(1,2,3,4,5,6)P_6$  o  $IP_6$ ). Sin embargo, en el caso de levaduras, sólo la acción de dos quinasas es necesaria para la síntesis de  $IP_6$  (**Figura 1.5**) (Irvine y Schell, 2001; Hatch y York, 2010).



En las sucesivas secciones se tratarán en detalle las diferentes familias y subfamilias de InsP Ks.

### 1.4.2 Inositol fosfato kinasas

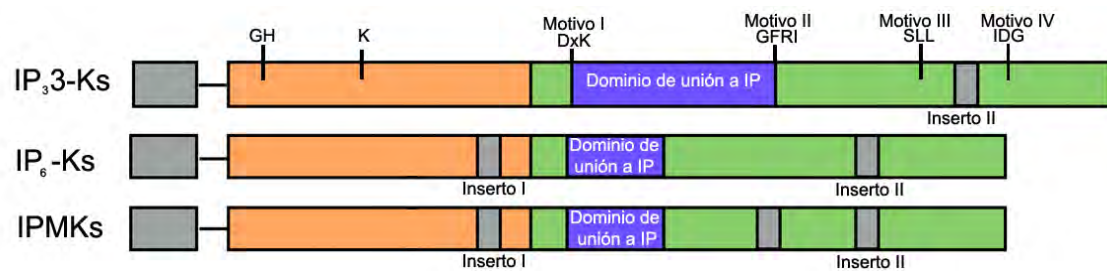
Las inositol fosfato kinasas son las enzimas encargadas de la síntesis de los diferentes inositol fosfatos. Existen diferentes familias de InsP Ks, las cuales presentan un dominio catalítico con actividad kinasa, presentando algunas otros dominios con diferentes funciones, principalmente reguladoras. Según la homología de secuencia y predicción de estructura secundaria de su dominio kinasa, las InsP Ks se han clasificado en (González *et al.*, 2004):

- **Familia I:** inositol polifosfato kinasas (IPKs): formada por inositol trifosfato 3 kinasas (IP<sub>3</sub> 3-Ks) (Choi *et al.*, 1990), inositol polifosfato multikinasa (IPMK) (Saiardi *et al.*, 1999) e inositol hexakisfosfato kinasas (IP<sub>6</sub>-Ks) (Saiardi *et al.*, 1999; Schell *et al.*, 1999). Estas enzimas presentan un plegamiento parcialmente conservado con el de las proteínas kinasas.
- **Familia II:** inositol (1,3,4) trisfosfato 5/6 kinasas (IP<sub>3</sub> 5/6-Ks), las cuales presentan un plegamiento de tipo ATP/GRASP (Wilson y Majerus, 1996).
- **Familia III:** inositol pentakisfosfato 2 kinasas (IP<sub>5</sub> 2-Ks) (York *et al.*, 1999), cuyo plegamiento se desconoce.

A continuación se tratará cada una de ellas en mayor detalle.

#### 1.4.2.1 Familia I. Inositol polifosfato kinasas

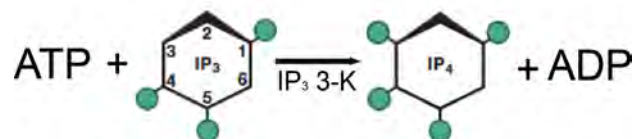
Las inositol polifosfato kinasas conforman la familia más extensa y mejor caracterizada de todas las inositol kinasas. Está compuesta por tres subfamilias: la subfamilia de las inositol trifosfato 3 kinasas (IP<sub>3</sub> 3-Ks), la subfamilia de las inositol polifosfato multikinastas (IPMKs) y la subfamilia de las inositol hexakisfosfato kinasas (IP<sub>6</sub>-Ks). Éstas presentan una serie de características comunes y también diferencias, las cuales se pudieron analizar una vez conocida una estructura para esta familia (González *et al.*, 2004; Miller y Hurley, 2004). Estas se resumen en la **Figura 1.6**.



**Figura 1.6. Organización general de la familia de las IPKs.** Los cuadrados grises indican regiones variables entre las diferentes enzimas. Las regiones en naranja corresponden al lóbulo N-terminal de las diferentes enzimas, mientras que el lóbulo C-terminal está representado en verde. Dentro del lóbulo C, se encuentra la zona de unión a IP (azul).

#### 1.4.2.1.1 Inositol trifosfato 3 kinasas (IP<sub>3</sub> 3-Ks)

La IP<sub>3</sub> 3-K es la inositol kinasa más estudiada, debido a su papel dual como punto de inicio en el metabolismo de los inositol fosfatos por un lado, y como punto clave en la regulación de la homeostasis celular del calcio por otro. Esta enzima es extremadamente específica tanto en el sustrato que une (IP<sub>3</sub>), como en el producto que sintetiza (inositol 1,3,4,5-tetraquisfosfato; IP<sub>4</sub>) (Irvine *et al.*, 1986), actuando como señal de finalización para la liberación de Ca<sup>2+</sup> desde el retículo endoplasmático (Pattni y Banting, 2004) (**Figura 1.5** y **Figura 1.7**). Además, el IP<sub>3</sub> también puede desfosforilarse mediante la acción de diferentes inositol polifosfato 5 fosfatasas (INPP5As) produciéndose inositol 1,4-bisfosfato (IP<sub>2</sub>), un inositol inactivo que la célula emplea como sustrato para la síntesis de PtdI(4,5)P<sub>2</sub> (Irvine y Schell, 2001).



**Figura 1.7. Reacción catalizada por la IP<sub>3</sub> 3-K.**

La IP<sub>3</sub> 3-K está presente en numerosos metazoos (Xia y Yang, 2005). Los mamíferos poseen tres isoformas de esta enzima, siendo cada una diferente al resto en cuanto a patrón de expresión, masa molecular y mecanismos de regulación. De humanos se han clonado y caracterizado las tres isoformas (Takazawa *et al.*, 1991; Takazawa *et al.*, 1991; Dewaste *et al.*, 2000). La isoforma A (51 kDa) se expresa principalmente en el prosencéfalo, en las células de Purkinje y en los testículos (Vanweyenberg *et al.*, 1995). La isoforma B, con una masa molecular de 100 kDa, se encuentra principalmente en el sistema inmunológico (linfocitos y neutrófilos) (Vanweyenberg *et al.*, 1995). Por



último, la isoforma C, de 75 kDa es ubicua, siendo su expresión particularmente activa en el músculo esquelético (Dewaste *et al.*, 2000).

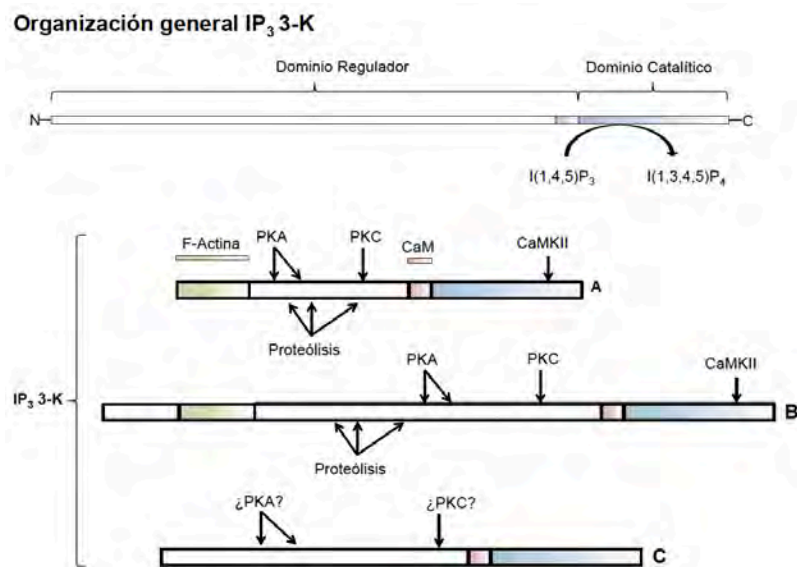
El producto de esta kinasa, el  $I(1,3,4,5)P_4$ , puede actuar como mensajero intracelular, aunque aún no está claro cuales son sus receptores (Schell, 2010). El  $I(1,3,4,5)P_4$  está implicado también en los denominados fenómenos de “memoria molecular”. En estos, una primera estimulación de la célula causaría un incremento del  $I(1,3,4,5)P_4$ . Un segundo estímulo a continuación del primero causaría un aumento del  $I(1,4,5)P_3$ , cuya vida media estaría prolongada (de ahí el nombre de “memoria”) debido a la presencia de  $I(1,3,4,5)P_4$  del estímulo anterior, el cual bloquearía la desfosforilación del  $I(1,4,5)P_3$  por parte de la INPP5A (Hermosura *et al.*, 2000; Irvine, 2001).

Las tres isoformas comparten una organización común. Presentan un dominio catalítico C-terminal, que es el más conservado tanto de unos organismos a otros como entre las diferentes isoformas (80% de identidad en la secuencia de aminoácidos de las isoformas A y B de rata). Además, cada isoforma posee una región N-terminal de longitud variable, la cual está destinada a la regulación de la actividad de la proteína mediante la unión de  $Ca^{2+}/CaM$  y a la localización subcelular de cada isoforma (Pattni y Banting, 2004; Xia y Yang, 2005) (**Figura 1.8**).

Las tres isoformas humanas de la  $IP_3$  3-K están reguladas por CaM, aunque en diferente grado (**Figura 1.8**). En el caso de la isoforma A humana, la unión de la CaM induce un aumento de la actividad del dominio kinasa de entre 2-3 veces por encima del nivel basal. Este aumento es mayor en el caso de la isoforma B (7-8 veces), y menor en la isoforma C (1.5 veces) (Woodring y Garrison, 1997; Dewaste *et al.*, 2000; Dewaste *et al.*, 2003). El CBD se encuentra, en los tres casos, unido a la zona N-terminal del dominio catalítico. La región que contiene las características típicas de un CBD ha sido localizada entre los aminoácidos 158 y 189 (en el caso de la isoforma A de rata), identificándose uno de los triptófanos de este fragmento como fundamental para la interacción (Takazawa y Erneux, 1991; Erneux *et al.*, 1993). Los alineamientos de secuencia indican que el motivo no se ajusta a ninguno de los tres tipos más comunes (Rhoads y Friedberg, 1997), aunque en el caso de la isoforma B se ha propuesto un motivo de tipo 1-10 (Chamberlain *et al.*, 2005). No hay muchos más datos acerca de esta interacción, ni se ha caracterizado a nivel estructural, desconociéndose como  $Ca^{2+}/CaM$  regula la actividad de  $IP_3$  3-K.

Además de esta regulación común mediante CaM, existen otros mecanismos de regulación de la actividad y la localización subcelular, específicos para cada isoforma (**Figura 1.8**). Las diferentes isoformas de  $IP_3$ -3K presentan sitios de fosforilación para 3 enzimas diferentes: CamKII, PKA y PKC (Pattni y Banting, 2004; Schell, 2010). La CamKII y la PKA son capaces de fosforilar las isoformas A y B, aumentando la  $V_{max}$

entre 8-10 veces (fosforilación por CamKII) (Communi *et al.*, 1997; Communi *et al.*, 1999) o 2 veces (fosforilación por PKA) (Sim *et al.*, 1990; Woodring y Garrison, 1997). La PKC es capaz de fosforilar las 3 isoformas, describiéndose tanto inhibición (Sim *et al.*, 1990; Woodring y Garrison, 1997) como activación de la actividad enzimática (Communi *et al.*, 1999).



**Figura 1.8. Organización general de la familia de las IP<sub>3</sub> 3-Ks.** Las regiones verdes indican las zonas de interacción con F-actina, las zonas rojas la región de interacción con CaM y la zona azul indica el dominio catalítico.

La localización subcelular de estas enzimas también se encuentra regulada. En el caso de la isoforma A humana, está principalmente asociada a F-actina en las espinas dendríticas (zona en la cual existe una gran acumulación de CaMKII). Los primeros 66 residuos de la enzima albergan una secuencia de 33 aminoácidos de longitud en la que se encuentra una región rica en prolinas, inmediatamente seguida de una hélice  $\alpha$ . Esta región es la encargada de la localización de esta isoforma (Schell *et al.*, 2001; Johnson y Schell, 2009). En el caso de la isoforma B, se ha encontrado también una secuencia para asociación a F-actina (Brehm *et al.*, 2004). En el caso de la isoforma C, se ha localizado una secuencia de exportación nuclear (NES) en el extremo N-terminal de la proteína, la cual sería responsable de la exportación de esta isoforma entre el núcleo y el citoplasma (Nalaskowski *et al.*, 2003). En las tres isoformas las secuencias N-terminal albergan también motivos PEST implicados en la localización subcelular (Pattni y Banting, 2004). Estos motivos, habituales en las proteínas reguladas por CaM, pueden ser el blanco de la acción de determinadas proteasas activadas por calcio, como por ejemplo las calpaínas (Ono y Sorimachi, 2012). La acción de las calpaínas tendría como

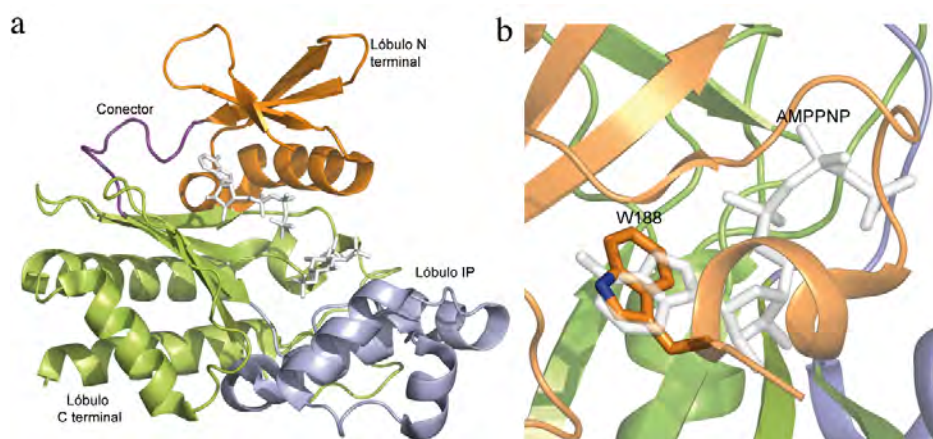
consecuencia la liberación del fragmento catalítico de la IP<sub>3</sub> 3-K al citosol (**Figura 1.8**) (Pattni *et al.*, 2003).

Hasta la fecha, el dominio catalítico es la única región de la proteína caracterizada estructuralmente. En el año 2004 se logró resolver la estructura del dominio kinasa de la isoforma A, tanto en su forma libre así como en complejo con sus sustratos o análogos (IP<sub>3</sub>/AMPPNP –un análogo no hidrolizable del ATP-) o productos (IP<sub>4</sub>/ADP) (González *et al.*, 2004; Miller y Hurley, 2004). Poco después se caracterizó la estructura del dominio kinasa de la isoforma B (Chamberlain *et al.*, 2005). Por último, se hizo pública la estructura de la isoforma C (código PDB 2a98). La estructura de la IP<sub>3</sub> 3-K se considera el prototipo para el resto de las inositol polifosfato kinasas. La estructura tridimensional del dominio kinasa de la enzima está dividida en 3 subdominios o lóbulos (**Figura 1.9a**). El primero, un subdominio  $\alpha+\beta$ , localizado en la región N-terminal (lóbulo N-terminal), consta de una lámina  $\beta$  de 4 hebras y una hélice  $\alpha$ . El dominio puede presentar una segunda hélice  $\alpha$  ( $\alpha 0N$ ) en el extremo N-terminal. Esta pequeña hélice, localizada entre los residuos 187-196 sólo puede observarse en ausencia de sustratos o productos estando desordenada en presencia de éstos. Un segundo subdominio  $\alpha+\beta$  (lóbulo C-terminal) se compone de una lámina  $\beta$  estabilizada mediante interacciones con 3 hélices  $\alpha$ . Los lóbulos N y C están unidos por un conector, y el nucleótido (ATP) se encuentra situado entre los lóbulos N y C, presentando unión con residuos de ambos. El tercer lóbulo (de unión a inositol) está exclusivamente formado por hélices  $\alpha$  y se encuentra insertado en el lóbulo C, encargándose de la unión del sustrato inositido (lóbulo IP) (**Figura 1.9a**).

Prácticamente todos los residuos que participan en la unión de I(1,4,5)P<sub>3</sub> e I(1,3,4,5)P<sub>4</sub> provienen del lóbulo de unión al inositido. En su mayoría son residuos básicos que conforman una amplia red de interacciones en forma de bolsillo, con el sustrato/producto. La magnitud y disposición del lóbulo de unión a IP explica dos hechos: por un lado la enorme especificidad de esta proteína, y por otro lado el hecho de que la enzima sea capaz de discriminar inositoles solubles en favor de inositoles lipídicos, como por ejemplo el PtdIns(4,5)P<sub>2</sub> (González *et al.*, 2004; Miller y Hurley, 2004).

La unión a ATP se realiza mediante un bolsillo hidrofóbico localizado en la interfase entre las láminas  $\beta$  de los lóbulos N y C, estando cerrado por uno de sus extremos por residuos pertenecientes al conector. En ausencia de nucleótido, uno de los residuos cercanos a la hélice  $\alpha 0N$  (W188) ocupa el bolsillo de unión a ATP. Esto podría representar un mecanismo de autoinhibición de la enzima, estando a su vez relacionado con la regulación de la enzima por parte de Ca<sup>2+</sup>/CaM (González *et al.*, 2004; Miller y Hurley, 2004) (**Figura 1.9b**).

El dominio kinasa de la  $\text{IP}_3$  3-K presenta ciertas características conservadas con las proteínas kinasas. Por ejemplo, el plegamiento general del lóbulo N-terminal está conservado, así como el reconocimiento de ATP entre ambos lóbulos. Otros elementos conservados son un bucle rico en glicinas implicado en la unión de los fosfatos del ATP (el denominado bucle G en proteínas kinasas) y un puente salino dentro del lóbulo N-terminal, determinante de la conformación activa de una proteína kinasa (Taylor *et al.*, 1999). Por otro lado existen diferencias notables, principalmente en el carácter totalmente helicoidal del dominio C-terminal en el caso de las proteínas kinasas frente a un dominio con una extensa lámina  $\beta$ , como es el caso de la  $\text{IP}_3$  3-K (González *et al.*, 2004; Miller y Hurley, 2004). Por ultimo, las  $\text{IP}_3$  3-Ks presentan un subdominio adicional: el lóbulo IP.

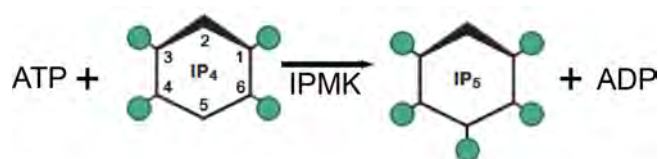


**Figura 1.9. Estructura de  $\text{IP}_3$  3-K.** a) Estructura general de  $\text{IP}_3$  3-K. El lóbulo N-terminal aparece en naranja, el C-terminal en verde y el lóbulo IP en azul claro. El conector entre los lóbulos N y C-terminal aparece en color violeta. Los ligandos  $\text{I}(1,4,5)\text{P}_3$  aparecen en color blanco. b) Detalle del sitio de unión a nucleótido. La figura muestra la superposición del complejo  $\text{IP}_3$  3-K/ $\text{I}(1,4,5)\text{P}_3$ /AMPPNP (en blanco se representa el AMPPNP perteneciente a este complejo) con la estructura de la forma vacía, empleando el código de colores que en a). Puede observarse cómo en la estructura que no contiene ligando el residuo W188 ocupa el sitio de unión a nucleótido.

#### 1.4.2.1.2 Inositol polifosfato multikinasa

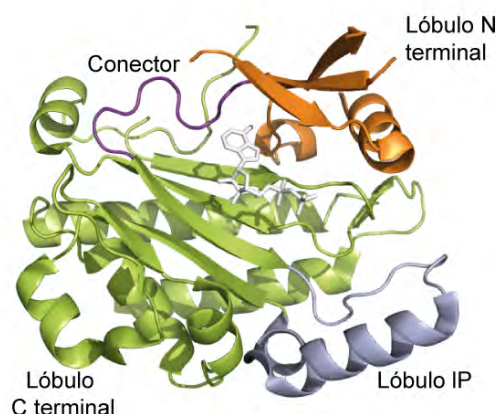
La Inositol Polifosfato Multikinasa (IPMK) es el segundo miembro de la familia de las Inositol Polifosfato Kinasas. Esta proteína ha sido clonada y caracterizada en múltiples organismos (Saiardi *et al.*, 1999; Odom *et al.*, 2000; Nalaskowski *et al.*, 2002). Al contrario que la  $\text{IP}_3$  3-K, la IPMK puede unir varios sustratos así como fosforilar diferentes posiciones del anillo de inositol. Concretamente, se ha observado actividad *in vitro* sobre las posiciones 3, 5 y 6, predominando una u otra actividad según el organismo (Shears, 2004) y (Figura 1.5 y Figura 1.10). En el caso de mamíferos su

actividad principal es la fosforilación del  $I(1,3,4,6)P_4$  en su posición 5, generando  $I(1,3,4,5,6)P_5$  (Shears, 2004; Holmes y Jogl, 2006) (**Figura 1.10**). En los organismos donde la  $IP_3$  3-K no se encuentra presente, esta enzima actúa primero como 3-kinasa, generando  $I(1,3,4,5)P_4$  y sintetizando posteriormente  $I(1,3,4,5,6)P_5$  (Odom *et al.*, 2000; Saiardi *et al.*, 2000). La IPMK cumple diferentes funciones en la célula. Recientemente se ha identificado como uno de los cofactores de mTOR (Kim *et al.*, 2011), demostrándose también que la IPMK es capaz de unirse y regular la actividad de la kinasa AMPK (Bang *et al.*, 2012), así como la proteína p53 (Xu *et al.*, 2013).



**Figura 1.10. Reacción catalizada por la IPMK.** Se muestra la reacción mayoritaria en mamíferos.

La estructura de la IPMK de *Saccharomyces cerevisiae* fue obtenida en el año 2006 (Holmes y Jogl, 2006) (**Figura 1.11**), determinándose recientemente la estructura de la IPMK de *A. thaliana* (Endo-Streeter *et al.*, 2012). En ambos casos, el plegamiento y organización es muy similar al descrito previamente para la  $IP_3$  3-K, a pesar de la escasa homología en secuencia que muestran estas dos familias de proteínas. La única diferencia notable reside en el dominio de unión a inositol, el cual es significativamente menor en las IPMKs, estando compuesto por una única hélice  $\alpha$ . El sitio de unión al inositido es mucho menos constreñido que el de  $IP_3$  3-K, lo cual explica por que estas enzimas son menos específicas; pueden albergar distintos sustratos y unirlos en distintas orientaciones permitiéndose la fosforilación del mismo anillo en distintas posiciones.

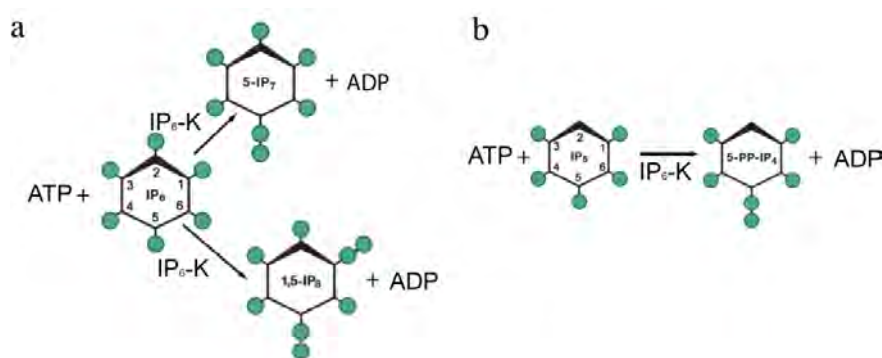


**Figura 1.11. Estructura de IPMK.** El código de colores es el mismo que en la **Figura 1.9**.



## 1.4.2.1.3 Inositol hexakisfosfato kinasa

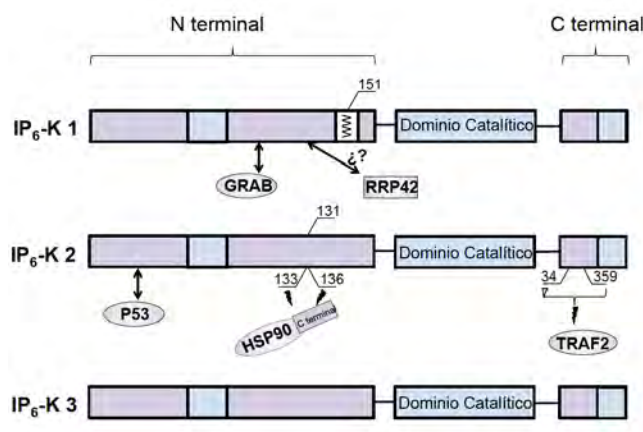
La inositol hexakisfosfato kinasa es la enzima que convierte el  $\text{IP}_6$  en  $5\text{-PP-IP}_7$  (anillo de inositol con fosfatos en las posiciones 1, 2, 3, 4 y 6 y un grupo pirofosfato en la posición 5) o en  $1, 5\text{-PP-IP}_8$  (anillo de inositol con fosfatos en las posiciones 2, 3, 4 y 6 y un grupo pirofosfato en las posiciones 1 y 5, respectivamente) (Saiardi, 2012) (**Figura 1.5** y **Figura 1.12a**). Esta enzima también puede utilizar como sustrato el  $\text{I}(1,3,4,5,6)\text{P}_5$ , pero con una afinidad mucho más reducida (Mulugu *et al.*, 2007), convirtiéndolo en  $5\text{-PP-IP}_4$  (un anillo de inositol con fosfatos en las posiciones 1, 3, 4 y 6 y un grupo pirofosfato en la posición 5) (**Figura 1.5** y **Figura 1.12b**). Los inositoles pirofosfatados están siendo muy estudiados en la actualidad, determinándose su participación en procesos de apoptosis (Morrison *et al.*, 2001; Nagata *et al.*, 2005), regulación del ciclo celular (Lee *et al.*, 2007), recombinación de DNA (Luo *et al.*, 2002), control de la longitud de los telómeros (Saiardi *et al.*, 2005) o quimiotaxis (Luo *et al.*, 2003) entre otros (Bennett *et al.*, 2006). Además, se ha demostrado que son capaces de transferir (a una serina previamente fosforilada) el fosfato  $\beta$  de su grupo pirofosfato sin necesidad de que ninguna proteína kinasa medie la transferencia (Saiardi *et al.*, 2004).



**Figura 1.12. Reacciones catalizadas por la IP<sub>6</sub>-K.** a) Reacciones empleando IP<sub>6</sub> como sustrato. b) Reacción empleando IP<sub>5</sub> como sustrato.

Existen tres isoformas diferentes de la enzima en humanos (**Figura 1.13**). La IP<sub>6</sub>-K 1 codifica para una proteína de 50 kDa, la cual se expresa principalmente en cerebro y testículos (Saiardi *et al.*, 1999). La delección de esta proteína en ratones *knock out* provoca una reducción en los niveles de insulina plasmática, sugiriendo que esta isoforma podría ser importante en procesos relacionados con la secreción de la insulina (Bhandari *et al.*, 2008; Chakraborty *et al.*, 2010). La enzima es capaz de asociarse mediante su extremo N-terminal a un factor intercambiador de nucleótidos de guanina denominado GRAB regulando con esta interacción la exocitosis en ciertos tipos celulares (Luo *et al.*, 2001). El N-terminal podría ser responsable también de la unión de

la proteína RRP42, una proteína implicada en el procesado del mRNA, aunque aún no se ha demostrado inequívocamente la interacción (Rual *et al.*, 2005).



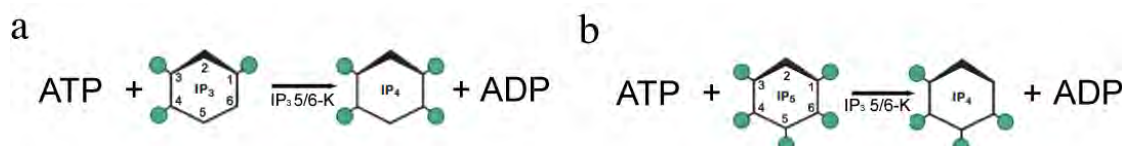
**Figura 1.13. Organización general de las IP<sub>6</sub>-Ks.** Las regiones azules indican las zonas más conservadas entre las isoformas, mientras que las zonas moradas indican las regiones más variables.

La segunda isoforma descrita de la IP<sub>6</sub>-K es la isoforma 2 (IP<sub>6</sub>-K 2) (**Figura 1.13**). Esta es una proteína de 49 kDa (Saiardi *et al.*, 1999; Schell *et al.*, 1999) que se encuentra distribuida en múltiples tejidos (Saiardi *et al.*, 1999). La isoforma 2 es bastante similar a las otras dos, sin embargo posee un N-terminal diferenciado, así como un inserto de aproximadamente 40 aminoácidos hacia el extremo C-terminal, los cuales son responsables de la unión a diferentes proteínas reguladoras (Barker *et al.*, 2009). La secuencia de la región C-terminal podría servirle a esta isoforma para interaccionar con la proteína TRAF2 (Morrison *et al.*, 2007). La región N-terminal podría ser responsable de la interacción de esta proteína con la chaperona Hsp90. La unión de Hsp90 bloquea la actividad kinasa de la enzima, probablemente bloqueando la entrada del IP<sub>6</sub> al sitio activo (Chakraborty *et al.*, 2008). La proteína también es capaz de unirse a p53 empleando su extremo N-terminal (Koldobskiy *et al.*, 2010). Estas interacciones (ambas específicas de la isoforma 2), junto con el papel proapoptótico de los inositol pirofosfatos, ha convertido recientemente a la IP<sub>6</sub>-K 2 en una de las enzimas claves para comprender los procesos de apoptosis (Koldobskiy y Snyder, 2011). La caracterización de un ratón *knock out* para esta enzima, el cual muestra una mayor predisposición a tumores, refuerza esta idea (Morrison *et al.*, 2009).

En cuanto a la isoforma 3, es la menos caracterizada (**Figura 1.13**). Se encuentra principalmente localizada en neuronas y es significativamente más ácida que las otras isoformas, así como más pequeña (46 kDa). Sus características cinéticas son muy similares a las de la isoforma 1, pero por el momento no se ha caracterizado su unión a ninguna proteína reguladora (Saiardi *et al.*, 2001).

### 1.4.2.2 Familia II. Inositol (1,3,4) trisfosfato 5/6 kinasas (IP<sub>3</sub> 5/6-Ks)

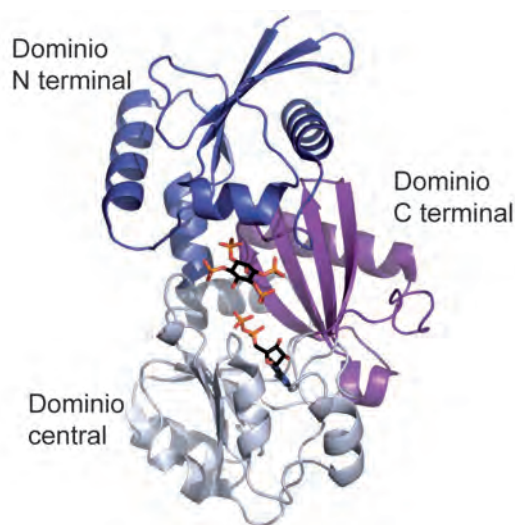
La familia II de Inositol Kinasas está formada por un único miembro: la Inositol (1,3,4) trisfosfato 5/6 kinasas (IP<sub>3</sub> 5/6-Ks). En mamíferos, esta enzima fue caracterizada en un primer momento como la encargada de fosforilar las posiciones 5 y 6 del anillo de inositol. Puede emplear como sustrato el I(1,3,4)P<sub>3</sub> (Yang y Shears, 2000) (**Figura 1.10** y **Figura 1.14a**). Este compuesto es producido por la acción de la fosfatasa INPP5 sobre el I(1,3,4,5)P<sub>4</sub> (el cual es el producto de la IP<sub>3</sub> 3-K). No obstante se ha observado que los productos liberados por la enzima humana son el I(1,3,4,5,6)P<sub>5</sub> y el I(3,4,5,6)P<sub>4</sub> (**Figura 1.5** y **Figura 1.14b**) lo cual implica que la enzima debe poseer actividad tanto kinasa como fosfatasa (Ho *et al.*, 2002). El I(3,4,5,6)P<sub>4</sub> posee una gran relevancia fisiológica, pues es un inhibidor de la conductancia de los canales de Cl<sup>-</sup> regulados por Ca<sup>2+</sup> (Vajanaphanich *et al.*, 1994). La estructura cristalográfica de la enzima humana (Chamberlain *et al.*, 2007) permitió conocer cómo están integrados los mecanismos kinasa y fosfatasa en una misma enzima: la enzima emplea un sistema en el que el inositol pentakisfosfato transfiere su fosfato en posición 1 a una molécula de ADP, liberándose I(3,4,5,6)P<sub>4</sub>. La molécula de ATP generada queda bloqueada en el centro activo, al cual a continuación se incorpora una molécula de I(1,3,4)P<sub>3</sub>, que es fosforilada en su posición 5 o 6 (Chamberlain *et al.*, 2007).



**Figura 1.14. Reacciones catalizadas por la IP<sub>3</sub> 5/6-K.** a) reacción empleando I(1,3,4)P<sub>3</sub> como sustrato. b) reacción empleando IP<sub>5</sub> como sustrato.

Hasta la fecha, dos estructuras de esta enzima han sido determinadas: las correspondientes a la enzima de *Entamoeba histolytica* (Miller *et al.*, 2005) y la anteriormente nombrada de *Homo sapiens* (Chamberlain *et al.*, 2007). Ambas estructuras presentan una organización similar (**Figura 1.15**). La enzima presenta un plegamiento de tipo ATP Grasp con tres dominios: un dominio N-terminal, un dominio central y un dominio C-terminal. El sitio de unión a ATP se encuentra localizado entre las láminas  $\beta$  del dominio central y del dominio C-terminal, mientras que los residuos que coordinan el inositol provienen fundamentalmente del dominio C-terminal (Miller *et al.*, 2005; Chamberlain *et al.*, 2007).

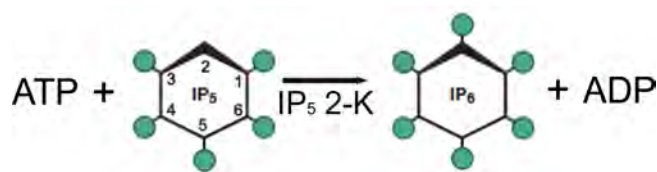




**Figura 1.15. Estructura de IP<sub>3</sub> 5/6-Ks.** El dominio N-terminal se representa en azul, el dominio central en blanco y el dominio C-terminal en morado. Los ligandos se representan en negro y naranja.

### 1.4.2.3 Familia III. Inositol pentakisfosfato 2 kinasas (IP<sub>5</sub> 2-Ks)

La tercera familia de inositol kinasas está formada por la inositol 1,3,4,5,6-pentaquisfosfato kinasas, o IP<sub>5</sub> 2-K. Esta enzima es capaz de fosforilar el 2-OH del anillo de *myo*-inositol, en concreto del inositol (1,3,4,5,6) pentakisfosfato (IP<sub>5</sub>), catalizando la síntesis de inositol (1,2,3,4,5,6) hexakisfosfato, (IP<sub>6</sub> o ácido fítico) (**Figura 1.5** y **Figura 1.16**) (Stephens *et al.*, 1991). Esta es la única IP K capaz de fosforilar un grupo OH en conformación axial, actuando el resto de IP Ks siempre sobre enlaces en posición ecuatorial. La enzima ha sido caracterizada de diferentes fuentes, aunque posee poca homología entre los diferentes organismos (11% de identidad desde levaduras a mamíferos) (Phillippy *et al.*, 1994; Verbsky *et al.*, 2002; Sweetman *et al.*, 2006). Es destacable el hecho de que no existe homología evidente con el resto de IP Ks, conformando por tanto una familia independiente dentro de las Inositol Kinasas. No obstante, algunos estudios proponen una homología muy limitada con éstas, así como con proteínas kinasas (Cheek *et al.*, 2002; Cheek *et al.*, 2005). La IP<sub>5</sub> 2-K es una enzima clave en el desarrollo de mamíferos. De hecho, la delección del gen codificante para la enzima (*Ipkl*) en ratones, tiene efectos letales para los embriones (Verbsky *et al.*, 2005). También se ha observado que la enzima está implicada en el establecimiento de la asimetría en el desarrollo de mamíferos (Sarmah *et al.*, 2005). Recientemente se ha observado que es un componente estructural del nucleolo, teniendo interacciones funcionales con las proteínas CK2, TCOF1 y UBF, las cuales están implicadas en la regulación del proceso de síntesis de RNA (Brehm *et al.*, 2012).



**Figura 1.16. Reacción catalizada por la IP<sub>5</sub> 2-K.**

El sustrato universal de esta enzima es, como se ha mencionado, el I(1,3,4,5,6)P<sub>5</sub>, aunque también puede catalizar la fosforilación de otros sustratos en determinados organismos (por ejemplo I(1,4,5,6)P<sub>4</sub> e I(3,4,5,6)P<sub>4</sub> en plantas (Sun *et al.*, 2007)). El *myo*-I(1, 3, 4, 5, 6)P<sub>5</sub> fue aislado originalmente de eritrocitos de aves (Johnson y Tate, 1969), siendo el isómero de IP<sub>5</sub> más frecuente (Stephens *et al.*, 1991). El IP<sub>5</sub> es un punto central en el metabolismo de inositol fosfatos, pues pese a la gran diversidad de rutas metabólicas existentes en los diferentes organismos, todas ellas acaban desembocando en la síntesis de IP<sub>5</sub> (Irvine y Schell, 2001). En distintas especies de animales con eritrocitos nucleados, este compuesto cumple la función de disminuir la afinidad de la hemoglobina por el oxígeno, sustituyendo en esta función al 2-bifosfoglicerato (Irvine y Schell, 2001). También se ha observado su participación en apoptosis (Maffucci *et al.*, 2005). Además, es uno de los inositoles capaz de unirse a dominios de tipo plectrina, de forma que podría regular la interacción (mediante competición) entre los dominios PH de diferentes proteínas y los fosfoinosítidos (Komander *et al.*, 2004). Es también un potente inhibidor de los canales de Ca<sup>2+</sup> regulados por I(1,4,5)P<sub>3</sub> (Lu *et al.*, 1996).

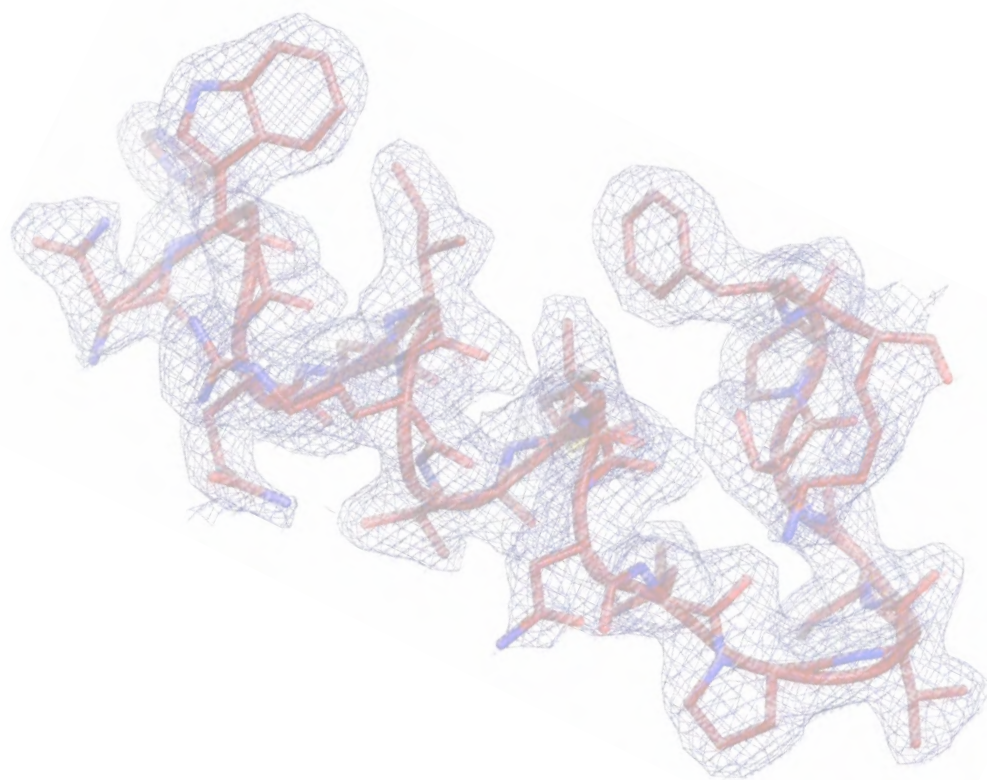
Por otro lado, el producto de reacción, IP<sub>6</sub>, es el inositol fosfato más abundante en la célula y uno de los más estudiados (Shi *et al.*, 2006). Este compuesto presenta múltiples funciones, por ejemplo, se ha demostrado su participación en procesos de maduración y exportación del mRNA. Para ello una de las proteínas implicadas en el proceso, Gle1 une IP<sub>6</sub>, activando el complejo una proteína llamada Dbp5, la cual participa en la exportación del mRNA hacia el citoplasma (Montpetit *et al.*, 2011). El IP<sub>6</sub> también participa en procesos de reparación del DNA, en concreto en el proceso de recombinación no homóloga (*Non Homologous End Joining*) (Cheung *et al.*, 2008). También se ha observado la participación del inositol hexakisfosfato en procesos relacionados con la regulación de la estructura de la cromatina (Shen *et al.*, 2003). Otro de los modos en los que el IP<sub>6</sub> ejerce su acción es el de actuar como cofactor de ciertas proteínas. Dos ejemplos destacados de ello son la enzima humana ADAR2 (Macbeth *et al.*, 2005), implicada en la edición de RNA, y la ubiquitina ligasa TIR1 presente en *A. thaliana* (Tan *et al.*, 2007).

El IP<sub>6</sub> también juega un papel crucial en plantas. Por ejemplo, participa activamente en la resistencia a patógenos (Murphy *et al.*, 2008). Además, en estos organismos, el

ácido fítico cumple otros papeles, principalmente en su desarrollo, pues es la mayor reserva de fósforo en la semilla (alrededor del 75% del fosfato total en la semilla). En las semillas, el  $\text{IP}_6$  se encuentra en forma de sal (fitato) acompañada con diferentes iones, como por ejemplo  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$ ,  $\text{Mn}^{2+}$  o  $\text{Zn}^{2+}$ . El fitato se acumula en ciertas zonas de la semilla (como por ejemplo la capa de aleurona) en forma de inclusiones denominadas globoides (Raboy y Bowen, 2006). Por otro lado, el gran poder quelante del  $\text{IP}_6$ , unido a su abundancia en las semillas, conlleva una serie de efectos en el ámbito de la salud y el medioambiente. Por un lado, en países cuya dieta se basa en la ingesta de cereales existen deficiencias de ciertos iones ( $\text{Zn}^{2+}$ ,  $\text{Fe}^{2+}$ ) en la población, actuando el  $\text{IP}_6$  como antinutriente (Hurrell, 2003). Sin embargo, la capacidad de quelar metales también posee una aplicación positiva, pues podría emplearse como agente de desintoxicación para envenenamientos por ciertos metales pesados (Cebrián *et al.*, 2007). Finalmente, los animales monogástricos no son capaces de digerir este compuesto, lo que provoca que al alimentarse de piensos que contienen fitato en altas cantidades se excrete en forma de sales que contribuyen a los procesos de eutrofización de las aguas. Por estos motivos, existe un gran interés en la obtención de semillas con bajos niveles en ácido fítico (Raboy, 2001).

Finalmente, como se señaló previamente, el  $\text{IP}_6$  es el precursor empleado por la  $\text{IP}_6\text{-K}$  para producir inositoles pirofosforilados con importantes funciones en apoptosis y modificación postraduccional de proteínas (Saiardi, 2012).

A pesar del papel fundamental tanto de esta enzima como de su sustrato y producto, y de su gran interés a nivel médico y biotecnológico, no se conoce ningún tipo de información estructural sobre la misma, ni se han podido identificar aspectos fundamentales de su función como son los residuos implicados en la actividad enzimática, en la unión a los ligandos, o sus mecanismos de regulación.



## 2 OBJETIVOS

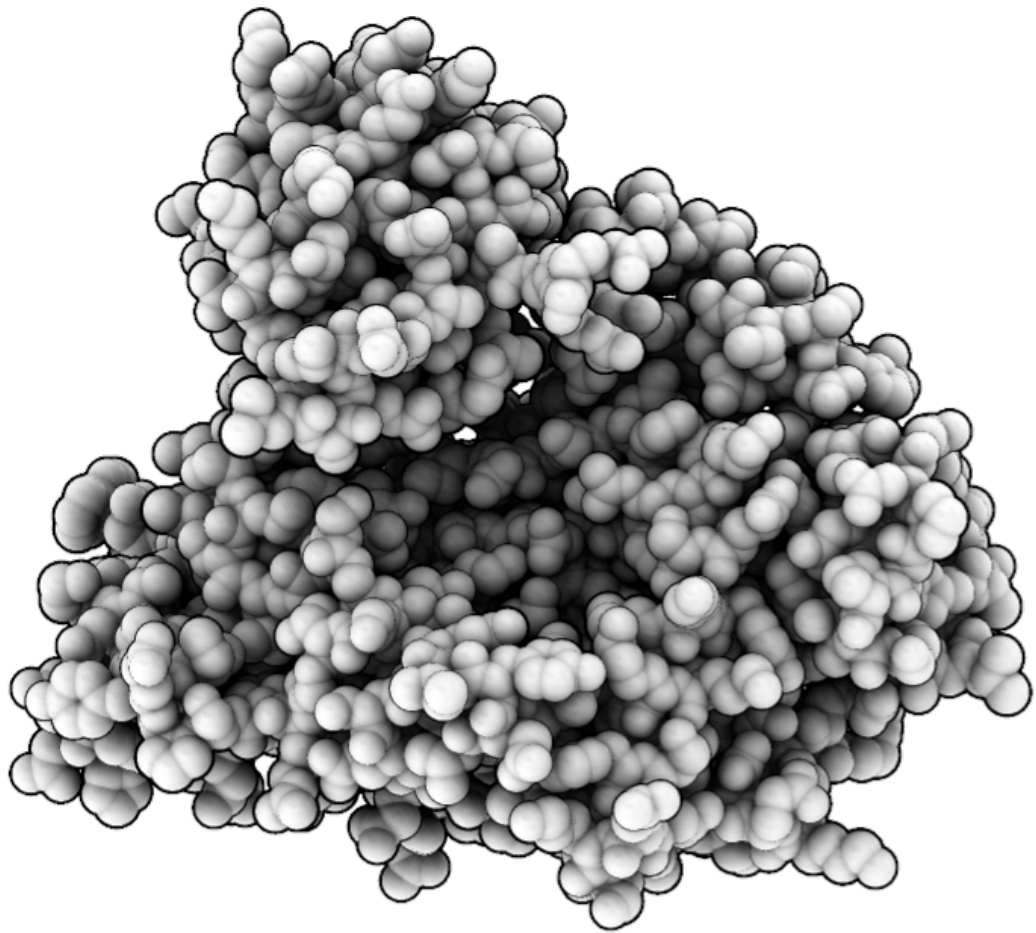


Como se ha expuesto en el apartado anterior, en los últimos años se ha hecho un gran avance en la caracterización de las InsP Ks debido a su importancia clave en diferentes procesos fisiológicos. Pese a ello, todavía existe gran desconocimiento sobre ciertos aspectos de estas enzimas. Concretamente, hasta la fecha no se ha caracterizado aún ninguna estructura perteneciente a la familia de las  $IP_5$  2-Ks (familia III). Por otro lado, y pese a que es la IP K más estudiada, se desconocen los determinantes estructurales de la regulación de  $IP_3$  3-K (familia I) por la CaM.

Por lo tanto, los objetivos concretos de este trabajo son:

- Determinación de la estructura de la  $IP_5$  2-K y estudio de la relación estructura-función en esta enzima:
  1. Clonación, expresión, purificación, cristalización y resolución estructural de la  $IP_5$  2-K de *A. thaliana*.
  2. Análisis estructural y funcional de la  $IP_5$  2-K de *A. thaliana*. Elucidación de los determinantes estructurales que rigen la unión de sustratos ( $IP_5$ , ATP) y productos ( $IP_6$ , ADP).
  3. Caracterización de los cambios estructurales ocurridos durante el ciclo catalítico de  $IP_5$  2-K.
- Determinación de la estructura del complejo entre el dominio de unión a CaM de la  $IP_3$  3-K A humana (CBD  $IP_3$  3-K) y la CaM humana:
  1. Obtención del complejo entre CBD  $IP_3$  3-K y CaM.
  2. Cristalización y resolución estructural de dicho complejo.
  3. Análisis estructural del complejo. Elucidación de los determinantes estructurales que rigen la unión de CaM y CBD  $IP_3$  3-K.
  4. Elaboración de un modelo del complejo entre CaM e  $IP_3$  3-K, mediante la combinación de datos a alta y baja resolución.





## 3 MATERIALES Y MÉTODOS

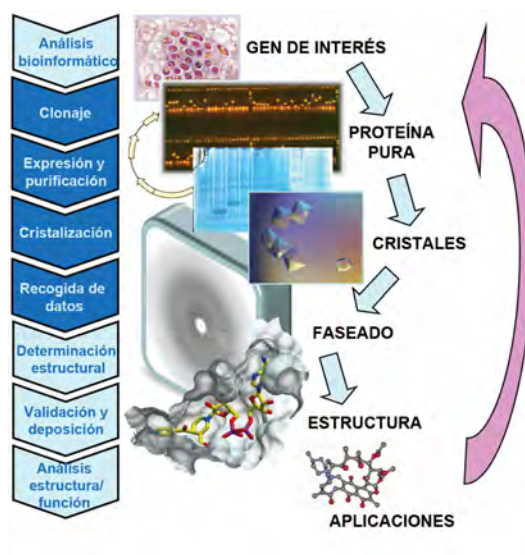




### 3.1 Determinación estructural de proteínas mediante cristalografía de rayos X

La cristalografía de rayos X es la técnica más potente para la determinación de la estructura tridimensional de moléculas a resolución atómica, con independencia de su tamaño. Esta técnica permite estudiar desde péptidos y proteínas de tamaño reducido a las más complejas máquinas macromoleculares.

En general, la resolución estructural de proteínas mediante cristalografía de rayos X conlleva sucesivas etapas (**Figura 3.1**). En los últimos años se han implementado técnicas de alto rendimiento en algunas de las etapas del proceso, como el uso de robots en el proceso de cristalización o las mejoras asociadas al desarrollo de sincrotrones de 3ª generación. También se han desarrollado enormemente los programas empleados, aumentando la automatización, gracias a la cual se han incrementado sustancialmente el número de estructuras de proteínas determinadas en los últimos años. Pese a esta automatización varias de sus etapas son consideradas aún “cuellos de botella”. Así, por ejemplo, la obtención de una muestra pura y homogénea, el propio proceso de cristalización de dicha muestra, o el cálculo de las fases requieren de un gran trabajo experimental para poder abordar con éxito el estudio estructural. Los detalles teóricos sobre cada fase del proceso completo pueden encontrarse en el Anexo 1 (Sección 7.1)



**Figura 3.1.** Etapas básicas de la cristalografía macromolecular de rayos X. Figura adaptada de Rupp (2009).

## 3.2 Metodología experimental

Tanto los materiales biológicos (bacterias, vectores empleados, enzimas y *kits* de biología molecular), así como otros reactivos y las secuencias de los diferentes oligonucleótidos utilizados en este trabajo se encuentran recogidos en el Anexo 2 (Sección 7.2).

### 3.2.1 IP<sub>5</sub> 2-K de *A. thaliana*

#### 3.2.1.1 Estructura cristalográfica de IP<sub>5</sub> 2-K en complejo con inositidos. Complejos con sustratos (IP<sub>5</sub>, IP<sub>5</sub>/AMPPNP e IP<sub>4</sub>/AMPPNP) y productos (IP<sub>6</sub> e IP<sub>6</sub>/ADP)

##### 3.2.1.1.1 Construcción de vectores de expresión

El gen *Ipkl* de *A. thaliana* fue amplificado mediante PCR empleando como molde el cDNA (At5G 42810 DNAC; (Sweetman *et al.*, 2006)), cedido por el Dr. Charles Brearley (*School of Biological Sciences, University of East Anglia*, Norwich, Reino Unido) y clonado en dos vectores de expresión diferentes: pOPTG, cedido por la Dra. Olga Perisic (*Laboratory of Molecular Biology, MRC, Cambridge*, Reino Unido) y pKLSL<sub>1</sub>, cedido por el Dr. José Miguel Mancheño (IQFR, CSIC, Madrid) (Angulo *et al.*, 2011). Para la PCR se emplearon los oligonucleótidos 1, 2, 3 y 4 respectivamente (la secuencia puede consultarse en el Anexo 2 (Sección 7.2)). La clonación se realizó de acuerdo a métodos ya publicados (Sambrook y Russell, 2001). Los plásmidos obtenidos contienen una secuencia codificante para la proteína glutatión S transferasa (GST) y para una lectina del hongo *Laetiporus sulphureus* (LSL) respectivamente, lo que permite obtener las correspondientes proteínas de fusión GST-IP<sub>5</sub> 2-K y LSL-IP<sub>5</sub> 2-K. Ambas etiquetas de fusión se encuentran situadas en posición N-terminal. Los plásmidos de expresión obtenidos fueron comprobados mediante secuenciación (Secugen S.L.).

##### 3.2.1.1.2 Expresión de proteína para estudios de cristalización

La construcción que contenía la proteína de fusión LSL-IP<sub>5</sub> 2-K se transformó en células de *E. coli* pertenecientes a la cepa Rosetta (DE3) pLysS (Novagen). A partir de las colonias obtenidas, se creció un precultivo a 310 K (37 °C) durante 6 horas, en medio 2TY (16 g/L bacto-triptona, 10 g/L extracto de levadura, 5 g/L de NaCl) al que se le añadió kanamicina hasta una concentración final de 50 µg/mL y cloranfenicol hasta 30 µg/mL. Se inocularon 10 mL de este cultivo en 900 mL de 2TY con

kanamicina y cloranfenicol y se dejaron crecer hasta una densidad óptica  $OD_{600} = 1$  a 310 K (37 °C). La inducción de la expresión se inició con la adición de IPTG hasta una concentración final de 0.3 mM, expresándose la proteína a 289 K (16 °C), durante 15 horas. El cultivo fue recogido mediante centrifugación durante 20 min, 1900 rcf, a 277 K (4 °C).

#### 3.2.1.1.3 Purificación

El cultivo anterior se resuspendió en tampón A (20 mM Tris-HCl pH 8.0, 0.15 M NaCl, 2 mM DTT), conteniendo éste una combinación de inhibidores de proteasas (Roche). A continuación, las células se lisaron mediante una prensa de French. El extracto celular se centrifugó a 277 K (4 °C) durante 45 minutos (38500 rcf) y se introdujo en una columna de *Sefarosa* CL-4B (GE Healthcare) equilibrada con el mismo tampón. La proteína fue eluída con tampón A suplementado con 0.2 M de lactosa. Seguidamente, la muestra fue diluida 1:3 con el fin de reducir su fuerza iónica y se introdujo en una columna de heparina HP de 5 mL (GE Healthcare) equilibrada con tampón B (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM DTT). La proteína fue eluída mediante un gradiente de NaCl entre 0-1 M empleando tampón C (20 mM Tris-HCl pH 8.0, 1 M NaCl, 2 mM DTT). La proteína de fusión fue digerida con la proteasa TEV (*Tobacco Etch Virus*) (1 mg de proteasa:80 mg de muestra) durante 12 horas a 277 K y en agitación suave. La muestra obtenida fue concentrada y aplicada en una columna 16/60 Superdex 200 (GE Healthcare) equilibrada en tampón D (20 mM Tris-HCl pH 8.0, 0.2 M NaCl, 2 mM DTT). El pico obtenido correspondiente a la proteína se aplicó en una columna de *Sefarosa* CL-4B (GE Healthcare) con el propósito de eliminar posibles contaminaciones por la proteína de fusión, empleando el propio tampón D para el equilibrado y la elución. La pureza de la muestra se comprobó mediante SDS-PAGE (Laemmli, 1970). La muestra se concentró hasta 10 mg/mL y se almacenó a 193 K.

#### 3.2.1.1.4 Cristalización

La búsqueda de las condiciones iniciales de cristalización para IP<sub>5</sub> 2-K se realizó mediante técnicas de alto rendimiento con un robot NanoDrop (Innovadyne Technologies Inc.), empleando las baterías de soluciones comerciales Index, SaltRx (Hampton Research), PACT Suite y JCSG+ Suite (Qiagen). En estos ensayos se empleó el método de la gota sentada en placas de 96 pocillos (microplacas Innovaplate SD-2; Innovadyne Technologies Inc.) manteniendo una temperatura controlada de 291 K (18 °C). Se mezclaron 250 nL de solución de proteína (10 mg/mL) con el mismo volumen de solución precipitante. La gota resultante se equilibró frente a un reservorio que contenía 60 µL de solución precipitante. El seguimiento de las condiciones de cristalización se realizó mediante una lupa de 40x (Nikon Corp.). Se realizaron ensayos duplicados empleando muestras incubadas con 2 mM de IP<sub>6</sub> (Sigma Aldrich), y muestras sin este producto. En estos ensayos iniciales se obtuvieron cristales en

diferentes condiciones que contenían PEG 3350 y PEG 6000, pero siempre en presencia de  $IP_6$ . Las condiciones fueron optimizadas en placas de 24 pocillos (placas Cryschem, Hampton Research), empleando el método de difusión de vapor en gota sentada. Para ello se mezcló 1  $\mu$ L de la solución de proteína pura (10 mg/mL) con 1  $\mu$ L de la solución precipitante, equilibrándose las gotas frente a un reservorio conteniendo 0.5 mL de solución precipitante. Durante el proceso de optimización se usaron diferentes técnicas: ajuste de la concentración de precipitante y de pH alrededor de las condiciones iniciales, el uso de diferentes aditivos (*Additive Screen*, Hampton Research), variación de la proporción proteína/precipitante durante la cristalización, o el uso de la técnica de sembrado. Esta optimización se realizó de forma independiente para cada uno de los complejos ( $IP_6$ ,  $IP_6$ /ADP,  $IP_5$ ,  $IP_5$ /AMPPNP e  $I(3,4,5,6)P_4$ /AMPPNP), obtenidos por cocrystalización, incubando ligando y muestras a 277 K durante 10 minutos previamente a la cristalización. La incubación se realizó con 1 mM de  $IP_5$  (Alexis), 1 mM de  $IP_5$  junto con 2 mM de AMPPNP (Sigma Aldrich), 2 mM de  $IP_6$  (Sigma Aldrich) y 2 mM de  $IP_6$  junto con 2 mM de ADP (Sigma Aldrich). Los cristales de mejor calidad fueron aquellos que contenían 2 mM  $IP_6$  y 2 mM ADP y 1 mM de  $IP_5$  junto con 2 mM de AMPPNP. Los cristales de los diferentes complejos aparecieron empleando concentraciones de entre 18% y 22% (m/v) de PEG 3350.

#### 3.2.1.1.5 Obtención y procesamiento de datos

##### 3.2.1.1.5.1 Crioprotección y derivatización con átomos pesados

Como paso previo a la recogida de datos se probaron múltiples soluciones crioprotectoras. El mejor crioprotector consistió en el aumento progresivo de la concentración de precipitante (PEG 3350) en la propia gota donde se encontraba el cristal, hasta una concentración de 35% (m/v) de PEG 3350 y 0.1 M de bis-tris pH 5.9. Todos los cristales de los diferentes complejos fueron crioprotegidos mediante este método.

Para el faseado de la estructura los mejores cristales ( $IP_6$ /ADP) fueron sometidos a inmersión en soluciones crioprotectoras (35% (m/v) de PEG 3350, 0.1 M Bis-Tris pH 5.9) que además contenían diferentes compuestos con átomos pesados (Hampton Research). Los compuestos empleados puede consultarse en la **Tabla 3.1**. Los cristales se mantuvieron en presencia de la solución derivada de átomo pesado durante diferentes tiempos, para posteriormente ser criogenizados en nitrógeno líquido a 100 K.

**Tabla 3.1 Condiciones empleadas para la derivatización con átomos pesados de IP<sub>5</sub> 2-K.**

Derivado	Tiempo de inmersión	Resolución (Å)	Experimento
I3C (ácido triiodo benzeno 5-amino-2,4,6 triiodoisoftálico)	30 minutos	2.2	No se apreció señal
NaI+2-Iodo ADP	30 minutos	3	No se apreció señal
NaI+8-Iodo ADP	30 minutos	3	No se apreció señal
TMLA (trimetil acetato de plomo)	30 minutos	3.3	MAD $\lambda = 0.948 \text{ \AA}$ (pico) $\lambda = 0.951 \text{ \AA}$ (inflexión) $\lambda = 0.945 \text{ \AA}$ (remoto)
Acetato de plomo	25 minutos	3.2	MAD $\lambda = 0.948 \text{ \AA}$ (pico) $\lambda = 0.950 \text{ \AA}$ (inflexión) $\lambda = 0.945 \text{ \AA}$ (remoto)
Acetato de samario	15 minutos	>3.5	SAD $\lambda = 1.541 \text{ \AA}$
Sulfato de gadolinio	15 minutos	>3.5	SAD $\lambda = 1.541 \text{ \AA}$

### 3.2.1.1.5.2 Recogida y procesamiento de datos

La recogida de datos de los diferentes complejos se realizó en todos los casos en el ESRF (Grenoble, Francia). Los distintos parámetros de la recogida de datos de cada complejo puede consultarse en la **Tabla 3.2**.

**Tabla 3.2 Parámetros de la recogida de datos de los diferentes complejos de IP<sub>5</sub> 2-K.**

	IP <sub>6</sub> /ADP/Pb <sup>2+</sup>	IP <sub>6</sub>	IP <sub>6</sub> /ADP	IP <sub>5</sub> /AMPPNP	IP <sub>5</sub>	IP <sub>4</sub> /AMPPNP	IP <sub>6</sub> /ADP (1 mol./u.a.)
<b>Parámetros de recogida</b>							
Número de imágenes	200	200	200	180	180	180	200
$\Delta\phi$ (°/imagen)	1	1	1	1	1	0.5	1
Línea de luz	ID23-1 (ESRF)	ID14-4 (ESRF)	ID14-1 (ESRF)	ID14-4 (ESRF)	ID23-2 (ESRF)	ID14-1 (ESRF)	ID14-1 (ESRF)
Longitud de onda (Å)	$\lambda_1 = 0.9478$ $\lambda_2 = 0.9503$ $\lambda_3 = 0.9442$	0.9330	0.9334	0.9395	0.9538	0.9334	0.9334
Exposición (s)	1	1	1	1	1	1	1
Detector	ADSC Q315R	ADSC Q315R	ADSC Q210	ADSC Q315R	Mar Mosaic225	ADSC Q210	ADSC Q210
Temperatura (K)	100	100	100	100	100	100	100

El procesamiento de todos los datos se realizó con el programa Mosfilm (Leslie, 2006), empleando para el escalado Scala (Evans, 2006), perteneciente al paquete de programas CCP4 (Winn *et al.*, 2011). El cálculo de los módulos de los factores de estructura fue realizado con el programa Truncate (French y Wilson, 1978), perteneciente al mismo paquete.

#### 3.2.1.1.6 Resolución estructural y construcción de los modelos

##### 3.2.1.1.6.1 Resolución estructural y refinamiento del complejo IP<sub>5</sub> 2-K IP<sub>6</sub>/ADP/Pb<sup>2+</sup>

La estructura de IP<sub>5</sub> 2-K fue resuelta a partir de un derivado obtenido con acetato de plomo (**Tabla 3.2**). Para la localización de los átomos pesados, se empleó el programa SHELXD (Sheldrick, 2008). Se encontraron 4 átomos de plomo, cuyas posiciones fueron refinadas y utilizadas para el cálculo de fases iniciales con el programa autoSHARP (Vonnrhein *et al.*, 2007). Posteriormente, y con este mismo programa, se realizó un procedimiento de mejora de fases mediante aplanamiento de solvente (para el cual se consideró un contenido de solvente de 43.2%). Empleando la información experimental, se construyó un primer modelo parcial que contenía en torno al 70% de los residuos existentes en la unidad asimétrica empleando el programa Buccaneer (Cowtan, 2008), también perteneciente al paquete CCP4. El resto del modelo fue construido manualmente, alternando ciclos de refinamiento con Refmac5 (Murshudov *et al.*, 2011) y ciclos de modelado manual con el programa O (Jones *et al.*, 1991).

##### 3.2.1.1.6.2 Resolución estructural y refinamiento de los complejos IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP, IP<sub>5</sub> 2-K /IP<sub>5</sub>/AMPPNP, IP<sub>5</sub> 2-K /IP<sub>6</sub>, IP<sub>5</sub> 2-K /IP<sub>5</sub> e IP<sub>5</sub> 2-K /IP<sub>4</sub>/AMPPNP

Para la resolución estructural de estos complejos, dado el isomorfismo entre los cristales, se utilizó el método de síntesis de diferencias de Fourier. Como modelo inicial se utilizó el modelo que contenía IP<sub>6</sub>/ADP, comenzándose por un refinamiento de tipo cuerpo rígido. Se eliminaron del modelo inicial todos los elementos que no eran cadena proteica. Los mapas obtenidos fueron de calidad suficiente para introducir los ligandos correspondientes. En el caso de los dos complejo no isomorfos la resolución estructural se llevó a cabo mediante Reemplazamiento Molecular (ver Sección 4.1.3)

El refinamiento de los modelos resultantes para cada complejo se realizó mediante el programa Refmac5 (Murshudov *et al.*, 2011) y el programa Coot (Emsley *et al.*, 2010).

La validación de todos los modelos se realizó con el programa Procheck (Laskowski *et al.*, 1993). Los modelos fueron depositados en el Banco de Datos de Proteínas (PDB). Su código de acceso puede consultarse en el Anexo 2 (Sección 7.2.6).

### 3.2.1.1.7 Análisis estructural

Las estructuras y los resultados obtenidos fueron analizados mediante diferentes programas. Los picos de la función de autorotación fueron analizados con el programa POLARRFN, perteneciente al paquete de programas CCP4 (Winn *et al.*, 2011). El análisis de plegamiento y búsqueda de homólogos estructurales se realizó con el servidor DALI (Holm y Rosenstrom, 2010). Los alineamientos de secuencia estructurales, predicciones de estructura secundaria y comparación con homólogos se realizaron utilizando la información de los programas ClustalW (Larkin *et al.*, 2007), Psi-Blast (Altschul *et al.*, 1997) y JPred (Cole *et al.*, 2008), ajustándolos manualmente con el programa Seaview (Gouy *et al.*, 2010). El cálculo de distancias entre residuos y ligandos fue realizado con el programa Contact, perteneciente al paquete de programas CCP4 (Winn *et al.*, 2011) y LigPlot (Wallace *et al.*, 1995). Las superporsiciones entre los diferentes complejos obtenidos se realizaron con el programa Coot (Emsley *et al.*, 2010). Los análisis gráficos se realizaron con los programas Coot y PyMOL. Las imágenes fueron realizadas con los programas PyMOL (Schrödinger, 2010) y Ligplot (Wallace *et al.*, 1995).

### 3.2.1.2 Estructura cristalográfica de IP<sub>5</sub> 2-K en estado libre y en complejo con nucleótido (IP<sub>5</sub> 2-K /AMPPNP e IP<sub>5</sub> 2-K /ADP)

#### 3.2.1.2.1 Obtención del vector de expresión para el mutante IP<sub>5</sub> 2-K<sub>W129A</sub>

El mutante W129A (IP<sub>5</sub> 2-K<sub>W129A</sub>), fue obtenido mediante mutagénesis dirigida (Papworth *et al.*, 1996), empleando para ello los oligonucleótidos 31 y 32 que aparecen en el Anexo 2. Como DNA molde se usó el plásmido que contenía la proteína nativa LSL-IP<sub>5</sub> 2-K. La introducción de la mutación fue comprobada por secuenciación del DNA (Secugen S.L.)

#### 3.2.1.2.2 Expresión y purificación de IP<sub>5</sub> 2-K<sub>W129A</sub> para estudios de cristalización

La IP<sub>5</sub> 2-K<sub>W129A</sub> fue expresada y purificada siguiendo el mismo protocolo que la proteína nativa (Apartado 3.2.1.1.2).



#### 3.2.1.2.3 Cristalización de IP<sub>5</sub> 2-K (IP<sub>5</sub> 2-K<sub>W129A</sub>)

##### 3.2.1.2.3.1 IP<sub>5</sub> 2-K<sub>W129A</sub> en estado libre

La búsqueda de las condiciones iniciales de cristalización se llevó a cabo con el procedimiento explicado en la Sección 3.2.1.1.4. Tras 2 semanas, sólo una de las 576 condiciones probadas produjo cristales (0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.5, 30% (m/v) PEG 4000). La condición fue optimizada en cajas de 24 pocillos (placas Cryschem, Hampton Research) mediante el método de gota sentada, mezclando 1 µL de la solución de proteína pura (10 mg/mL) con 1 µL de la solución precipitante. Durante la optimización se variaron las concentraciones de precipitante, así como el pH, obteniendo cristales en forma de placas, no adecuados para su estudio por difracción de rayos X (0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.0, 32–34% (m/v) PEG 4000). Una vez ajustados estos parámetros, se realizaron experimentos de sembrado empleando la técnica de *streak seeding*, lo que finalmente permitió obtener cristales de una calidad aceptable para su estudio mediante difracción de rayos X. Estos cristales se obtuvieron mezclando 1 µL de proteína con 1 µL de solución precipitante compuesta de 0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.0, y 29% (m/v) PEG 4000.

##### 3.2.1.2.3.2 Complejos IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP e IP<sub>5</sub> 2-K<sub>W129A</sub>/ADP

Con el fin de obtener cristales de IP<sub>5</sub> 2-K<sub>W129A</sub> en complejo con los nucleótidos (AMPPNP o ADP), se llevó a cabo la misma aproximación que para la forma apo, salvo que la proteína se incubó previamente con 2 mM de ADP (Sigma Aldrich) o 2 mM de AMPPNP (Sigma Aldrich), respectivamente. De nuevo, de entre todas las condiciones iniciales solo a partir de una se obtuvieron cristales de ambos complejos (0.5 M de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M citrato sódico pH 5.6 y 1 M de LiSO<sub>4</sub>). El proceso de optimización de estos cristales se basó en la técnica de sembrado (*streak seeding*) y el ajuste de la concentración de precipitante. Los mejores cristales se obtuvieron, en ambos casos, con 0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M citrato sódico pH 5.6 y 0.8 M LiSO<sub>4</sub>.

#### 3.2.1.2.4 Recogida y procesamiento de datos

Los cristales correspondientes a la forma apo de IP<sub>5</sub> 2-K<sub>W129A</sub> fueron crioprottegidos aumentando la concentración del PEG hasta un 35% (m/v) (solución crioprotectora: 0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.0, 35% (m/v) PEG 4000). Los cristales de IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP e IP<sub>5</sub> 2-K<sub>W129A</sub>/ADP fueron crioprottegidos mediante su inmersión en la propia condición de cristalización suplementada con un 20% de glicerol. Los detalles sobre la recogida de datos de estas muestras pueden consultarse en la **Tabla 3.3**.

**Tabla 3.3** Parámetros de recogida de datos de IP<sub>5</sub> 2-K<sub>W129A</sub> y los complejos IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP y IP<sub>5</sub> 2-K<sub>W129A</sub>/ADP.

	IP <sub>5</sub> -2K <sub>W129A</sub>	IP <sub>5</sub> -2K <sub>W129A</sub> /AMPPNP	IP <sub>5</sub> -2K <sub>W129A</sub> /ADP
<b>Parámetros de recogida</b>			
Número de imágenes	180	180	180
$\Delta\phi$ (°/imagen)	1	1	1
Línea de luz	ID14-4 ESRF	ID23-2 ESRF	ID14-4 ESRF
Longitud de onda (Å)	0.9478	1.0053	0.9334
Exposición (s)	2	1	1
Detector	ADSC Q315R	ADSC Q315R	ADSC Q315R
Temperatura (K)	100	100	100

El procesamiento de datos fue realizado con los mismos procedimientos que los explicados en la Sección 3.2.1.1.5.

#### 3.2.1.2.5 Resolución estructural y construcción de los modelos

Todas las estructuras fueron en este caso resueltas mediante Reemplazamiento Molecular con el programa Phaser (McCoy *et al.*, 2007), incluido también en el paquete CCP4. Como modelo inicial se utilizó el modelo de IP<sub>5</sub> 2-K obtenido a mayor resolución (complejo IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP). Dado que no se encontraba solución de reemplazamiento pese a la casi total homología del modelo con la proteína en estudio, se dividió el modelo en dos conjuntos: N-terminal (aminoácidos 3-150), y C-terminal (aminoácidos 160-437), realizando una búsqueda secuencial de ambos (primero C y luego N-terminal), que permitió finalmente obtener una solución. El refinamiento matemático se realizó con Refmac5 (Murshudov *et al.*, 2011), alternándose éste con la reconstrucción manual del modelo utilizando Coot (Emsley *et al.*, 2010). La validación se realizó con el programa Procheck (Laskowski *et al.*, 1993). Todos los modelos obtenidos fueron depositados en el Banco de Datos de Proteínas (PDB) (Berman *et al.*, 2003). Su código de acceso puede consultarse en el Anexo 2 (Sección 7.2.6).

#### 3.2.1.2.6 Análisis estructural

Para realizar el cálculo de distancias interatómicas entre residuos y residuo-ligando se empleó el programa Contact (Winn *et al.*, 2011). Las diferentes estructuras obtenidas se superpusieron usando Coot (Emsley *et al.*, 2010). Los cambios en IP<sub>5</sub> 2-K entre las diferentes formas cristalizadas fueron analizados tanto de forma manual como de forma automática con el programa Dyndom (Hayward y Berendsen, 1998). Las imágenes fueron realizadas con el programa PyMOL (Schrödinger, 2010).

#### 3.2.1.3 Análisis bioquímico de IP<sub>5</sub> 2-K

Los análisis bioquímicos fueron realizados en colaboración con el laboratorio del Dr. Charles Alastair Brearley (*School of Biological Sciences, University of East Anglia, Norwich, Reino Unido*). En concreto, los ensayos realizados por su equipo fueron los ensayos de cuantificación de zinc en la muestra (Sección 3.2.1.3.2), los ensayos cinéticos y de actividad enzimática (Sección 3.2.1.3.3) y los ensayos de fluorescencia (Sección 3.2.1.3.4).

##### 3.2.1.3.1 Obtención de Mutantes de IP<sub>5</sub> 2-K

El plásmido con la proteína de fusión GST-IP<sub>5</sub> 2-K se utilizó como patrón para realizar un primer grupo de mutantes mediante mutagénesis dirigida, según el protocolo descrito en Papworth *et al.*, (1996). Para la realización de los diferentes mutantes se emplearon los oligonucleótidos 5-30, que aparecen en el Anexo 2 (Sección 7.2). Se realizaron los mutantes: R40V, E85A, R130I, K168A, K168N, K170S, N238A, C330S, C333S, H346, D368A, D407A y K411A.

La expresión se realizó en las mismas condiciones que la de la enzima nativa, salvo por la presencia de 100 µg/mL de ampicilina (en lugar de kanamicina). Posteriormente el cultivo se resuspendió en tampón B (20 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM DTT), conteniendo éste un cóctel de inhibidores de proteasas (Roche). A continuación, las células se lisaron mediante prensa de French. El extracto celular fue centrifugado a 277 K (4 °C) durante 45 minutos (6500 rcf) y este lisado fue cargado en una columna de heparina HP (GE Healthcare) que fue equilibrada con tampón B. La proteína fue eluída mediante un gradiente entre tampón B y C (20 mM Tris-HCl pH 8.0, 1 M NaCl, 2 mM DTT). La muestra se cargó a continuación en una columna de Glutación-Sefarosa HP (GE Healthcare) equilibrada en tampón E (50 mM Tris-HCl pH 7.5, 0.2 M NaCl y 2mM DTT). La etiqueta de fusión fue eliminada mediante incubación en columna con la proteasa TEV empleando 1 mg de proteasa por cada 80 mg de muestra durante 12 horas a 277 K (4 °C). La muestra final (IP<sub>5</sub> 2-K) fue eluída con tampón E. La pureza de la muestra fue comprobada mediante SDS-PAGE. La muestra final fue almacenada a 193 K (-80 °C) hasta su empleo en los ensayos de actividad y fluorescencia.

Para los mutantes C330S, C333S, H346 se empleó el mismo método, pero utilizando una resina de unión a GST glutatión Sefarosa 4B (GE Healthcare).

Un segundo grupo de mutantes fue producido empleando como molde el plásmido codificante para la enzima nativa LSL-IP<sub>5</sub> 2-K. Estos mutantes fueron R45A (oligonucleótidos 33 y 34), W129A/R130A (oligonucleótidos 35/36, para esta mutación se usó como molde el plásmido que ya contenía la mutación W129A), R130A (oligonucleótidos 37/38) y delección de S253/S55 ( $\Delta$ S253/S55; oligonucleótidos 39/40).

La secuencia de estos oligonucleótidos puede consultarse en el Anexo 2 (Sección 7.2). Los mutantes fueron preparados según Papworth *et al.* (1996). La introducción de las mutaciones fue comprobada mediante secuenciación (Secugen S.L.). Los diferentes mutantes fueron expresados y purificados de la misma forma descrita para la proteína nativa (LSL-IP<sub>5</sub> 2-K), salvo por la eliminación de la cromatografía de exclusión molecular durante el proceso de purificación. Para la eliminación de la etiqueta de fusión tras el corte fue empleada una columna Sefarosa CL-4B. Los mutantes se concentraron entre 0.5 mg/mL-1 mg/mL y se almacenaron a 193 K (-80 °C) hasta su uso.

En todos los casos, se obtuvo enzima nativa en las mismas condiciones para usar como referencia en los diferentes ensayos realizados.

#### 3.2.1.3.2 Cuantificación de zinc en la muestra

La cuantificación de la cantidad de zinc en la muestra se realizó mediante espectrometría de emisión atómica por plasma acoplado inductivamente. Los experimentos se realizaron a 293 K (20 °C) con la muestra en tampón 5 mM Tris-HCl, pH 7.3 en un espectrómetro Vista Pro ICP-OES (Varian Technologies).

#### 3.2.1.3.3 Ensayos cinéticos y de actividad enzimática

##### 3.2.1.3.3.1 Ensayos de actividad enzimática de IP<sub>5</sub> 2-K nativa y mutantes

Tanto para la enzima nativa como para los mutantes, la actividad se midió en una solución consistente en 10 mM de HEPES pH 7.5, 6 mM MgCl<sub>2</sub>, 2 mM EDTA y 2 mM DTT. Además cada muestra analizada contenía 0.2 mM de ATP, 0.1 mM de IP<sub>5</sub>, 3.7 kBq de [ $\gamma$ -<sup>33</sup>ATP] y 0.08  $\mu$ g de proteína, todo ello en un volumen final de 10  $\mu$ l. Las reacciones se realizaron a una temperatura de 310 K (37 °C) durante 5 minutos. La medida correspondiente a cada muestra se repitió por triplicado y el conjunto del experimento por duplicado. Para los mutantes con baja actividad se realizó un experimento adicional, realizando el ensayo con una concentración 10 veces mayor de proteína. Los productos de todos los ensayos fueron separados mediante cromatografía líquida de alta presión en fase reversa (*reverse-phase HPLC*). En ninguno de los casos se observó liberación de fosfato inorgánico durante la reacción (es decir, actividad fosfatasa).

Para los ensayos cinéticos, la actividad kinasa fue estudiada mediante la medida de la liberación de ADP, a una concentración fija de 0.1 mM de IP<sub>5</sub> y un rango de concentraciones de ATP de entre 0.01 y 0.2 mM. Los ensayos se realizaron mediante el mismo protocolo que los anteriores.

#### 3.2.1.3.3.2 Efecto del EDTA sobre la actividad de IP<sub>5</sub> 2-K

En paralelo se midió el efecto del tratamiento con EDTA sobre la actividad enzimática. En este caso, fue determinado en un ensayo de acoplamiento (Rowan *et al.*, 2009). La proteína fue preincubada durante 30 minutos a 273 K con 2 mM de DTT, 50 mM de NaCl y 50 mM de Tris-HCl pH 7.3 en el caso del ensayo control, y 2 mM de DTT, 50 mM de EDTA y 50 mM de Tris-HCl en el caso de la proteína tratada. Una vez realizado este tratamiento, las muestras se diluyeron hasta un volumen de 0.7 mL conteniendo 450 ng de proteína, 116  $\mu$ M-ATP, 11.6  $\mu$ M IP<sub>5</sub> y 10 mM MgCl<sub>2</sub>. La reacción se monitorizó a una longitud de onda de 340 nm en un espectrofotómetro Cary 50 UV/Visible (Varian Technologies). Cada experimento fue repetido un total de 3 veces.

#### 3.2.1.3.3.3 Ensayos de actividad enzimática de los mutantes W129A, R45A W129A/R130A R130A y $\Delta$ S253/S55

Los ensayos enzimáticos fueron realizados en un volumen de 0.2 mL cada uno, usando un lector de placas Spectra Max M5 (Molecular Devices). El ensayo se realizó en un tampón conteniendo HEPES 20 mM pH 7.5, 1 mM de MgCl<sub>2</sub>, 0.1 mM de ATP, 2 mM de P-enoilpiruvato, 0.15 mM de NADH, 7.5 unidades por mL de lactato deshidrogenasa, 15 unidades por mL de PK y 1.5 mM del inositol fosfato correspondiente. Las reacciones se iniciaron mediante la adición de la proteína (nativa o mutante) a la solución descrita. Los ensayos se llevaron a cabo durante 15 min a 298 K (25 °C). Los ensayos se realizaron con 5 experimentos independientes en cada ensayo y cada ensayo fue repetido 3 veces. La reacción se monitorizó a una longitud de onda de 340 nm en un espectrofotómetro Cary 50 UV/Visible (Varian Technologies).

#### 3.2.1.3.3.4 Ensayos cinéticos de las muestras IP<sub>5</sub> 2-K y IP<sub>5</sub> 2-K<sub>W129A</sub>

Los ensayos cinéticos fueron realizados con 100 ng de proteína nativa o del mutante W129A. Se realizaron 5 experimentos independientes en cada ensayo, que se repitieron 3 veces. Los valores de  $K_{m\ ap}$  y  $V_{max\ ap}$  fueron obtenidos mediante regresión no lineal por mínimos cuadrados en una representación de la  $V_{ap}$  frente a la concentración de sustrato. Los valores se ajustaron a una ecuación de Michaelis-Menten empleando GraFit 5.0 (Erithcus Software Ltd).

#### 3.2.1.3.4 Medidas de fluorescencia

##### 3.2.1.3.4.1 IP<sub>5</sub> 2-K nativa vs IP<sub>5</sub> 2-K tratada con EDTA

Para la realización de los ensayos 28 ng de proteína fueron diluídos en 1 mL de tampón compuesto por 2 mM de DTT, 50 mM de NaCl y 50 mM Tris-HCl pH 7.3. Para el ensayo se empleó una cubeta de cuarzo de 3x10 mm. El ensayo de medición de fluorescencia se realizó a 298 K (25 °C) con un espectrofotómetro de fluorescencia Cary

Eclipse (Varian Technologies). Para la excitación se usó una longitud de onda de 295 nm empleando un paso óptico de 5 nm. El EDTA fue añadido en incrementos sucesivos de concentración de 0.5 mM, entre concentraciones de 0 y 2.5 mM. Los sucesivos espectros fueron recogidos con un retraso de 5 minutos respecto al anterior.

#### 3.2.1.3.4.2 Unión de ligandos en IP<sub>5</sub> 2-K nativa vs IP<sub>5</sub> 2-K<sub>W129A</sub>

Los cambios en la fluorescencia intrínseca de IP<sub>5</sub> 2-K y del mutante W129A durante la unión a los sustratos fueron registrados a 298 K (25 °C) mediante un espectrofotómetro de fluorescencia Cary Eclipse (Varian Technologies). Para la realización del experimento se escogió una longitud onda de excitación de 295 nm, y se registró la emisión de fluorescencia entre 310 y 380 nm. Para realizar el ensayo la proteína se disolvió en un tampón HEPES 20 mM pH 7.5 y 1 mM MgCl<sub>2</sub> y fue situada en una cubeta de cuarzo de 10x3 mm, realizando sucesivas adiciones de IP<sub>5</sub> o AMPPNP según el caso. Los cambios en fluorescencia a 335 nm  $[(F_o - F)/F_o]$  fueron representados frente a la concentración de ligando. Para ello se realizó una regresión no lineal de mínimos cuadrados ajustando la curva a un modelo de un solo sitio de unión. Se empleó para ello el programa GraFit 5.0 (Erithcus Software Ltd).

#### 3.2.1.3.5 Ensayos mediante geles nativos

Los ensayos se realizaron tanto para la proteína nativa como para el mutante W129A. Para realizarlos, se empleó un método derivado de Laemmli (Laemmli, 1970). Para ello se modificó la composición de todos los tampones de tal forma que ni estos ni los geles de acrilamida incluyesen ningún tipo de agente desnaturizante ni reductor. El experimento se realizó a 277 K, fijando el voltaje de la electroforesis en 40 V, y dejando que se desarrollara durante 6 horas. El porcentaje de acrilamida del gel empleado fue 12%. Para las muestras incubadas con IP<sub>6</sub> se empleó una concentración de 2 mM de este compuesto. En total se emplearon 10 µg de proteína por ensayo. Ninguna de las muestras se hirvió antes de ser cargada en el gel.

#### 3.2.1.3.6 Ensayos de proteólisis limitada

Los ensayos se realizaron tanto para la proteína nativa como para el mutante W129A. El ensayo se realizó como está descrito en (Gosein *et al.*, 2012). Se emplearon geles del 12% incubándose las muestras con una proporción tripsina/muestra de 1:100 (m/m) (Sigma Aldrich Aldrich). La digestión se llevó a cabo durante 20 minutos a 277 K (4 °C). En el caso de las proteínas incubadas con producto se empleó 2 mM de IP<sub>6</sub>.

## 3.2.2 Complejo IP<sub>3</sub> 3-K A/CaM de *Homo sapiens*

### 3.2.2.1 Estructura mediante cristalografía de rayos X del complejo entre el dominio de unión a CaM (CBD) de IP<sub>3</sub> 3-K A y CaM

#### 3.2.2.1.1 Construcción de vectores de expresión del CBD de IP<sub>3</sub> 3-K A y CaM humanas

En este trabajo se han producido tres construcciones truncadas de IP<sub>3</sub> 3-K humana que contenían diferentes fragmentos de IP<sub>3</sub> 3-K, concretamente entre los aminoácidos 158-201 (IP<sub>3</sub> 3-K<sub>(158-201)</sub>), 158-189 (IP<sub>3</sub> 3-K<sub>(158-189)</sub>) y 158-183 (IP<sub>3</sub> 3-K<sub>(158-183)</sub>). Para la primera construcción se empleó como molde un plásmido que contenía la enzima IP<sub>3</sub>-3K A humana (aminoácidos 128-461) clonada en el vector pOPTG (la construcción fue cedida por la Dra. Olga Perisic, *Laboratory of Molecular Biology*, MRC, Cambridge, Reino Unido), empleando los oligonucleótidos número 41/42. Este fragmento fue digerido y ligado con el vector pKLSL<sub>t</sub> en los sitios correspondientes.

A partir de esta segunda construcción se llevó a cabo la obtención de las otras dos construcciones mediante mutagénesis dirigida (Papworth *et al.*, 1996). Se emplearon los oligonucleótidos número 43 y 44 (Anexo 2 (Sección 7.2)) con el fin de sustituir el aminoácido 190 por un codón de parada de forma que el fragmento proteico codificado contuviera los aminoácidos 158 a 189. Se realizó el mismo procedimiento para obtener el fragmento comprendido entre los aminoácidos 158 y 183 (oligonucleótidos 45-46) (Anexo 2 (Sección 7.2)).

En todos los casos la secuencia de las construcciones fue comprobada mediante secuenciación (Secugen S.L.). Las digestiones, ligaciones y amplificaciones de DNA se han realizado de acuerdo a los métodos descritos en (Sambrook y Russell, 2001). El plásmido de expresión para la CaM (CaM/pET14b<sub>m</sub>) fue cedido por la Dra. Olga Perisic (*Laboratory of Molecular Biology*, MRC, Cambridge, Reino Unido).



### 3.2.2.1.2 Expresión y purificación de proteína para estudios de cristalización

#### 3.2.2.1.2.1 Expresión y purificación de CaM humana

Para la expresión de CaM humana el plásmido CaM/pET14b<sub>m</sub> fue transformado en *E. coli* BL21 (DE3). El crecimiento se realizó empleando medio 2TY, a una temperatura de 298 K (16 °C) hasta que la DO<sub>600</sub> alcanzó 1.5, momento en el que se indujo la expresión de proteína mediante la adición al cultivo de IPTG, en una concentración final de 0.3 mM. Una vez realizada la inducción, el cultivo se mantuvo en agitación durante 8 horas a 310 K (37 °C). Transcurrido este tiempo, el cultivo fue recogido mediante centrifugación durante 20 min, 1900 rcf, a 277 K (4 °C).

La purificación se realizó mediante el método descrito por (Hayashi *et al.*, 1998). El cultivo fue resuspendido en tampón de lisis (50 mM Tris-HCl pH 7.5, 2 mM EDTA y 0.2 mM PMSF) y lisado mediante el uso de prensa de French. El lisado se clarificó mediante centrifugación a una velocidad de 38500 rcf durante 1 hora (a 277 K; 4 °C). Una vez recuperado, el lisado fue filtrado con filtros de 0.45 µm. Para posibilitar la unión de la proteína a la columna se añadió al lisado 5 mM de CaCl<sub>2</sub>. El lisado fue cargado en una columna de HP fenil-Sefarosa (GE Healthcare), equilibrada con tampón de equilibrado (50 mM Tris-HCl pH 7.5, 5 mM CaCl<sub>2</sub> y 100 mM NaCl). Una vez introducido el lisado, la columna fue lavada con tampón de lavado 1 (50 mM Tris-HCl pH 7.5, 0.1 mM CaCl<sub>2</sub> y 100 mM NaCl) y 2 (50 mM Tris-HCl pH 7.5, 0.1 mM CaCl<sub>2</sub> y 500 mM NaCl). Finalmente la proteína fue eluída mediante un gradiente entre tampón de lavado 2 y tampón de elución (50 mM Tris-HCl pH 7.5, 2.0 mM EGTA).

#### 3.2.2.1.2.2 Expresión y purificación de la IP<sub>3</sub>- 3K A humana (fragmentos IP<sub>3</sub> 3-K<sub>158-201</sub>, IP<sub>3</sub> 3-K<sub>158-189</sub> e IP<sub>3</sub> 3-K<sub>158-183</sub>)

La expresión de los tres fragmentos fue realizada con un protocolo similar. El plásmido se transformó en la cepa Rosetta (DE3) pLyss de *E. coli*. El cultivo creció en medio 2TY a 310 K hasta que alcanzó una DO<sub>600</sub> de 1.0. La expresión de la proteína fue inducida mediante la adición de 0.3 mM de IPTG (concentración final), realizándose la expresión 16 h a 289 K. Las células fueron recogidas a 3400 rcf (277 K; 4 °C) y almacenadas a 193 K (-80 °C).

Al comienzo de la purificación, las células fueron resuspendidas en tampón A (20 mM Tris-HCl pH 8.0, 0.15 M NaCl, 2 mM DTT), el cual contenía una mezcla comercial de inhibidores de proteasas (Roche). Seguidamente se realizó una lisis mediante prensa de French. El extracto celular se centrifugó a 277 K (4 °C) durante 60 minutos (38500 rcf). El sobrenadante de ésta fue filtrado con filtros de 0.45 µm y mezclado con resina de *Sefarosa* CL-4B (GE Healthcare), previamente lavada con el tampón A. El lisado fue incubado con la resina durante una hora en agitación a 277 K (4



°C). Seguidamente, el lisado se descartó y la resina fue lavada repetidas veces con tampón A. A continuación, se adicionó proteasa TEV (1 mg de proteasa:80 mg de muestra) dejando actuar dicha proteasa durante 12 horas a 277 K (4 °C) en agitación suave. Por último, el correspondiente fragmento de IP<sub>3</sub> 3-K A fue recuperado mediante lavados sucesivos de la resina con el tampón A.

#### 3.2.2.1.2.3 Formación y purificación de los complejos IP<sub>3</sub> 3-K<sub>158-201</sub>/CaM, IP<sub>3</sub> 3-K<sub>158-189</sub>/CaM y IP<sub>3</sub> 3-K<sub>158-183</sub>/CaM

Los complejos entre CaM y los CBD de IP<sub>3</sub> 3-K A se obtuvieron mezclando ambas proteínas en una proporción molar 1:3 (relación CaM/CBD) en presencia de 2 mM de CaCl<sub>2</sub>. Esta mezcla fue concentrada y cargada en una columna Hiload 16/600 Superdex 75, previamente equilibrada en tampón E (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 2 mM CaCl<sub>2</sub>). La elución del complejo se realizó con el mismo tampón. La adecuada formación y pureza de la muestra se analizó mediante SDS-PAGE. Las fracciones que contenían el complejo fueron reunidas, concentradas hasta 25 mg/mL y almacenadas a 193 K (-80 °C) hasta su empleo en los ensayos de cristalización.

#### 3.2.2.1.3 Cristalización del complejo IP<sub>3</sub> 3-K<sub>158-183</sub>/CaM

Con el fin de cristalizar el complejo entre el dominio de unión a CaM de IP<sub>3</sub> 3-K A y CaM se empleó la aproximación de alto rendimiento explicada en la Sección 3.2.1.1.4 con los tres complejos producidos.

Tras cuatro semanas de incubación el complejo IP<sub>3</sub> 3-K<sub>(158-183)</sub> produjo pequeños cristales con forma cúbica en una de las condiciones probadas (0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M acetato de sodio pH 4.6, 25% (m/v) de PEG 4000). Con el fin de obtener cristales de mayor tamaño se realizó un barrido fino del pH y concentración de PEG, empleando la técnica de gota sentada en placas Cryschem de 24 pocillos (Hampton Research). Gotas con proporción 1μL:1μL (proteína:precipitante) se equilibraron contra reservorios que contenían 0.5 mL de la condición precipitante. Cristales de un tamaño ligeramente superior se obtuvieron tras 8 semanas de incubación con la siguiente solución precipitante: 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M pH acetato de sodio pH 4.0 y 29% (m/v) de PEG 4000. El límite de difracción de estos cristales se encontraba entre 3.8-3.3 Å. Con el fin de mejorar la calidad del patrón de difracción, se emplearon diferentes técnicas, entre ellas la adición de aditivos y detergentes (Hampton Research) a la gota de cristalización. Para la realización de estos ensayos de cristalización se mezcló la solución de proteína (2.5 μL), la solución precipitante (2 μL) y un determinado aditivo o detergente comercial (0.5 μL). Se obtuvieron cristales muy similares a los previamente obtenidos mediante el uso de espermidina como aditivo a una concentración final en la gota de 10 mM. Con uno de los detergentes (óxido de n-undecil-N,N-dimetilamina) a una concentración final en la gota de 3.21 mM se obtuvieron también cristales, si bien con un hábito diferente a los obtenidos previamente.

#### 3.2.2.1.4 Recogida y procesamiento de datos

Todos los cristales fueron crioprotectados mediante el uso de la propia condición de cristalización (0.2 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.1 M pH acetato de sodio pH 4.0 y 29% (m/v) de PEG 4000) suplementada con un 20% de glicerol.

Los parámetros de recogida de datos de cada uno de ellos pueden observarse en la **Tabla 3.4**.

**Tabla 3.4** Parámetros de recogida de datos de CaM/IP<sub>3</sub> 3-K CBD.

	Hábito 1 (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 1 + espermidina (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 2 (CaM)
<b>Parámetros de recogida</b>			
Número de imágenes	360	174	90
$\Delta\phi$ (°/imagen)	1	1	1
Línea de luz	PXIII (SLS)	XALOC (ALBA)	XALOC (ALBA)
Longitud de onda (Å)	1.0000	0.9794	0.9794
Exposición (s)	1.5	1	1
Detector	PILATUS 2M	PILATUS 6M	PILATUS 6M
Temperatura (K)	120	100	100

El procesamiento de datos fue realizado mediante el mismo procedimiento que se explicó en la Sección 3.2.1.1.5.2.

#### 3.2.2.1.5 Resolución estructural y construcción de los modelos

Para la resolución estructural del complejo, se empleó el servidor Balbes (Long *et al.*, 2008) con el objetivo de encontrar un modelo inicial adecuado así como una primera solución de Reemplazamiento Molecular.

Empleando el conjunto de datos obtenido en el SLS (Tabla 3.4) se buscaron, con el servidor Balbes, modelos de partida que incluyeran únicamente las coordenadas de la CaM. Con uno de ellos (código PDB 2vay), se encontró una solución de Reemplazamiento Molecular (valor Q de 0.779, probabilidad de solución correcta del 99%). No obstante, los mapas de densidad calculados a partir de este modelo (3.8 Å) no permitieron el trazado sin ambigüedad del CBD y tener una visión detallada de la interacción. Gracias a la obtención de los cristales con espermidina se obtuvieron datos de mejor calidad permitiendo finalmente obtener el modelo completo (Tabla 4.2). Para ello se realizó un Reemplazamiento Molecular utilizando como modelo inicial el

obtenido anteriormente con Balbes y el programa Phaser (McCoy *et al.*, 2007). En este caso, los mapas obtenidos sí permitieron construir el fragmento del CBD de IP<sub>3</sub> 3-K.

También se obtuvo información estructural a partir de los cristales obtenidos en presencia de detergente. La estructura se resolvió mediante Reemplazamiento Molecular. Para ello, tras búsqueda en el PDB, se utilizó como modelo inicial una estructura cristalográfica que presentaba un grupo espacial y parámetros de celdilla muy similares a los cristales obtenidos (código PDB 2hqw) (Ataman *et al.*, 2007).

El ajuste manual de los modelos obtenidos se realizó con el programa Coot (Emsley *et al.*, 2010), realizándose el refinamiento matemático con el programa Refmac5 (Murshudov *et al.*, 2011). La validación de todos los modelos se realizó con el servidor MolProbity (Chen *et al.*, 2010).

#### 3.2.2.1.6 Análisis estructural

El cálculo de distancias entre residuos fue realizado con el programa Contact, perteneciente al paquete de programas de CCP4 (Winn *et al.*, 2011), mientras que el cálculo del área de superficies se realizó mediante el programa Areaimol (Lee y Richards, 1971) perteneciente al paquete de programas de CCP4. Las imágenes fueron realizadas con los programas PyMOL (Schrödinger, 2010) y UCSF Chimera (Pettersen *et al.*, 2004).

#### 3.2.2.2 Estructura mediante dispersión a bajo ángulo (SAXS) del complejo IP<sub>3</sub> 3-K/CaM

##### 3.2.2.2.1 Expresión y purificación del complejo IP<sub>3</sub> 3-K<sub>(128-461)</sub> para estudios de SAXS

###### 3.2.2.2.1.1 Expresión y purificación de CaM

La expresión y purificación de CaM se llevó a cabo del mismo modo que se explicó en la Sección 3.2.2.1.2.1.

###### 3.2.2.2.1.2 Expresión y purificación de IP<sub>3</sub> 3-K<sub>(128-461)</sub>

Para la obtención de IP<sub>3</sub> 3-K (128-461) se utilizó un plásmido que comprendía los residuos 128-461 de IP<sub>3</sub> 3-K humana insertados en el vector pOPTG (cedido por la Dra. Olga Perisic (*Laboratory of Molecular Biology*, MRC, Cambridge, Reino Unido). Con este plásmido, se transformaron células de *E. coli* pertenecientes a la cepa C41

(Novagen). A partir de las colonias obtenidas, se creció un precultivo a 310 K (37 °C) durante 6 horas, en medio 2TY (16 g/L bacto-triptona, 10/L g extracto de levadura, 5 g de NaCl) con una concentración de 100 µg/mL de ampicilina. Se inocularon 10 mL de este cultivo en 1000 mL de 2TY fresco con el antibiótico mencionado y se dejaron crecer hasta una densidad óptica  $OD_{600} = 1$  a 310 K (37 °C). La inducción de la expresión se inició con la adición de IPTG hasta una concentración final de 0.3 mM, expresándose la proteína a 289 K (16 °C), durante 15 horas. El cultivo fue recogido mediante centrifugación durante 20 min, 1900 rcf, a 277 K (4 °C). Las células fueron resuspendidas en tampón F (20 mM Tris-HCl pH 7.5 and 100 mM NaCl) y lisadas mediante prensa de French.

El lisado celular se centrifugó a 277 K (4 °C) durante 60 minutos (6500 rcf). El sobrenadante se filtró con filtros de 0.45 µm y posteriormente se mezcló con resina de *Sefarosa* CL-4B (GE Healthcare), previamente lavada con el tampón G (50 mM Tris-HCl pH 7.5 y 200 mM NaCl). El lisado fue incubado con la resina durante una hora en agitación a 277 K (4 °C). Seguidamente, el lisado se descartó y la resina fue lavada repetidas veces con tampón G. A continuación, se adicionó proteasa TEV (1 mg de proteasa:80 mg de muestra) dejando actuar dicha proteasa durante 12 horas a 277 K en agitación suave. Por último, el correspondiente fragmento de  $IP_3$  3-K<sub>(128-461)</sub> fue recuperado mediante lavados sucesivos de la resina con el tampón G. La  $IP_3$  3-K<sub>(128-461)</sub> obtenida fue introducida en una columna de heparina HiTrap (GE Healthcare), previamente equilibrada con tampón H (50 mM Tris-HCl pH 7.5 y 50 mM NaCl). La proteína fue eluída mediante un gradiente entre tampón H y tampón I (50 mM Tris-HCl pH 7.5 and 1 M NaCl). La proteína fue almacenada a 193 K (-80 °C) hasta su empleo en la formación del complejo  $IP_3$  3-K<sub>(128-461)</sub>/CaM.

#### 3.2.2.2.1.3 Purificación del complejo $IP_3$ 3-K<sub>(128-461)</sub>/CaM

La purificación del complejo  $IP_3$  3-K<sub>(128-461)</sub>/CaM se realizó siguiendo el mismo protocolo explicado para la purificación de los complejos  $IP_3$  3-K<sub>158-201</sub>/CaM,  $IP_3$  3-K<sub>158-189</sub>/CaM e  $IP_3$  3-K<sub>158-184</sub>/CaM para cristalografía. La única modificación consistió en que las fracciones que contenían el complejo fueron reunidas, concentradas a tres concentraciones diferentes (0.89, 5.02 y 13.56 mg/mL) y almacenadas a 193 K (-80 °C)

#### 3.2.2.2.2 Obtención de un modelo del complejo $IP_3$ 3-K<sub>(128-461)</sub>/CaM mediante SAXS

##### 3.2.2.2.2.1 Recogida y procesamiento de datos de SAXS

Los datos de difracción a bajo ángulo del complejo entre  $IP_3$  3-K/CaM fueron recogidos en la línea BM29 del ESRF (Grenoble, Francia), empleando una longitud de onda de 0.9919 Å. Previamente a la recogida de datos, las muestras fueron centrifugadas durante 10 minutos, siendo la parte inferior de la solución descartada con

el objetivo de eliminar posibles agregados de gran tamaño que pudieran interferir en las medidas. Para realizar las medidas, 40  $\mu\text{L}$  de una solución de proteína fueron introducidos en un capilar de cuarzo de 2 mm el cual se encontraba en un ambiente de vacío, empleándose un sistema automatizado de carga de muestras, compatible con tubos de PCR. La solución inicial de proteína se diluyó con su correspondiente tampón para preparar nueve soluciones de concentraciones conocidas (0.23, 0.40, 0.89, 2.08, 3.67, 5.02, 9.9, 13.56 y 14.83 mg/mL). A continuación, se recogieron de cada una de ellas 10 imágenes consecutivas, comenzando por la más diluida, con un tiempo de exposición de 2 segundos para cada una, que dieron lugar a 10 perfiles o curvas de dispersión. El detector empleado fue un Pilatus 1M (Dectris Ltd), el cual se encontraba a una distancia de la muestra 2.841 m. Todas las medidas se realizaron a 298 K (20 °C). Se recogieron también curvas de dispersión para el tampón (blanco) en el que se encontraba la proteína, tanto antes como después de la medida de cada muestra, para corregir las lecturas de fondo del detector. Puede verse un resumen de estos parámetros en la **Tabla 3.5**.

El procesamiento de los datos de dispersión se llevó a cabo utilizando el conjunto de programas del paquete ATSAS 2.3 (Petoukhov *et al.*, 2012). El análisis del perfil de dispersión experimental, la substracción de la dispersión del tampón, el escalado por la concentración y la combinación de las curvas tomadas a diferentes concentraciones, se realizaron con el programa PRIMUS (Konarev *et al.*, 2003) Puede verse un resumen de estos parámetros en la **Tabla 3.5**.

#### 3.2.2.2.2 Obtención de la envoltura y modelado del complejo $\text{IP}_3$ 3- $\text{K}_{(128-461)}/\text{CaM}$

Cada una de las diez exposiciones se comparó para asegurar su similitud y así descartar posibles daños por radiación. Los datos de SAXS para las diferentes concentraciones de proteína fueron evaluados usando los diagramas de Kratky, y explorados en busca de agregación usando los diagramas de Guinier. Las regiones de Guinier y el radio de giro ( $R_g$ ) se derivaron de la aproximación de Guinier  $I(q)=I(0)\exp(-q^2R_g^2/3)$ , donde  $qR_g < 1.3$ , usando la función AutoRg de Primus, donde  $q = 4\pi \sin q/\lambda$ . El examen de la serie de concentraciones no evidenció problemas de agregación o repulsión en los datos de alta concentración y por tanto mayor señal.

El programa GNOM se usó para el cálculo de la función de distribución pareada,  $P(R)$ , para cada una de las muestras, para determinar el tamaño máximo de partícula,  $D_{max}$ .

Las estructuras tridimensionales a baja resolución se obtuvieron mediante métodos *ab initio* de reconstrucción de forma (envoltorios moleculares), y fueron generadas por el programa DAMMIF (Franke y Svergun, 2009). Este programa crea un modelo de bolas (*dummy atom model*) que toma como referencia una esfera con un diámetro igual

a la distancia máxima de la molécula ( $D_{\max}$ ), rellenando con bolas un volumen correspondiente a la dispersión experimental de la muestra. Los diferentes modelos *ab initio*, fueron promediados mediante el programa DAMMAVER (Volkov y Svergun, 2003).

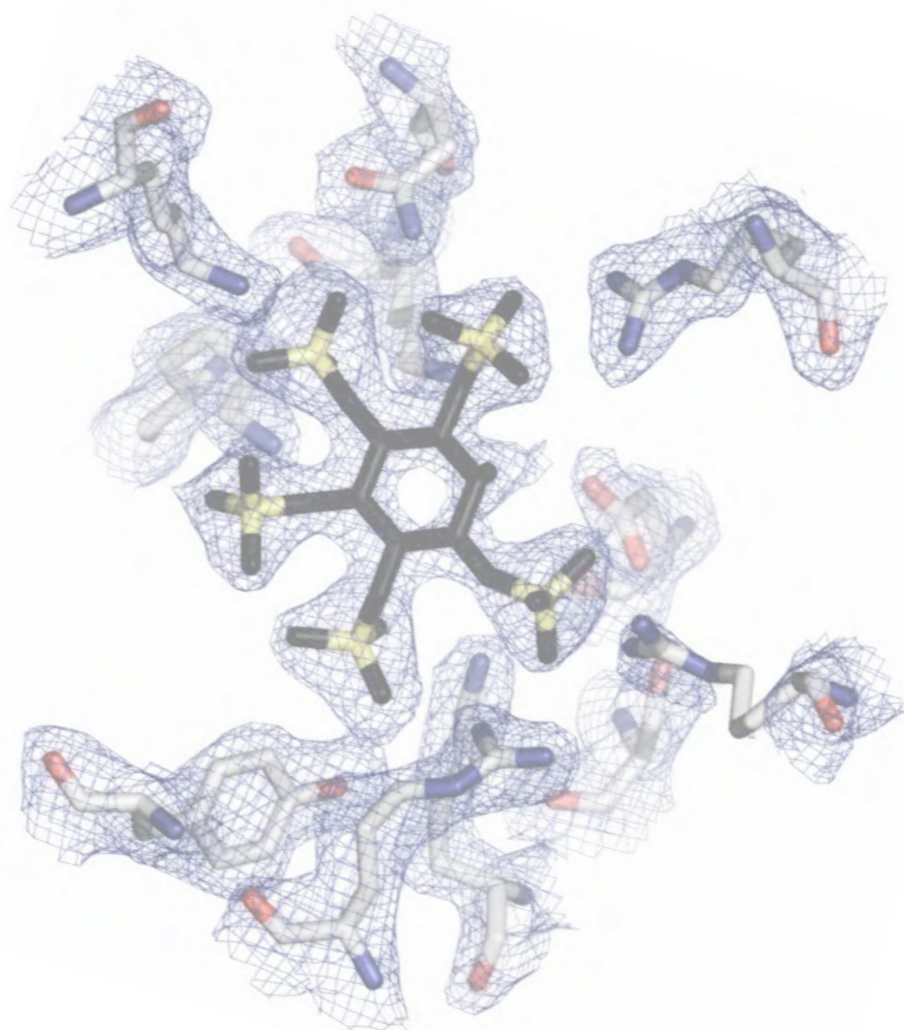
El programa CRY SOL (Svergun *et al.*, 1995) se usó para calcular la intensidad de dispersión (*scattering intensity*), a partir de los datos cristalográficos correspondientes a la estructura del dominio kinasa de IP<sub>3</sub> 3-K resuelta previamente (código pdb 1W2C; (González *et al.*, 2004)) y de la estructura del complejo entre IP<sub>3</sub> 3-K<sub>(158-184)</sub>/CaM, presentada en este trabajo. Usando el modelo/envoltorio *ab initio* obtenido por SAXS, se modeló el complejo IP<sub>3</sub> 3-K<sub>(128-461)</sub>/CaM. Para ello se empleó el programa Bunch (Petoukhov y Svergun, 2005), con el que se realizó un modelado de tipo cuerpo rígido. Las partes para las que no se contaba con una estructura cristalográfica (residuos 128-158 y residuos 184-189) de la IP<sub>3</sub> 3-K A, fueron modeladas por el propio programa mediante el añadido de cadenas de bolas que representaban la posición más probable de los residuos ausentes (*dummy residue model*).

**Tabla 3.5. Parámetros de recogida y procesado de datos de SAXS.**

Parámetros de recogida y procesado de datos	
<b>Recogida de datos</b>	
Línea	BM29 (ESRF)
Longitud de onda (Å)	0.9919
Rango $q$ (Å <sup>-1</sup> )	0.007-0.30
Tiempo de exposición (s)	2
Rango de concentración (mg ml <sup>-1</sup> )	0.23-14.83
Temperatura (K)	298
<b>Programas empleados</b>	
Reducción de datos y procesado de datos	PRIMUS
Calculo de modelos <i>ab initio</i>	DAMMIF
Validación y promediado de modelos <i>ab initio</i>	DAMMAVER
Modelado de cuerpo rígido/construcción de zonas ausentes	BUNCH
Cálculo de la intensidades del modelo	CRY SOL
Representación gráfica de resultados	PyMoL







## 4 RESULTADOS Y DISCUSIÓN

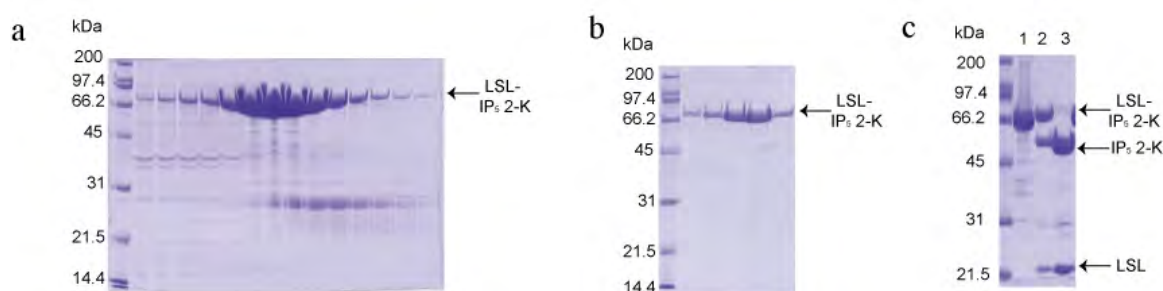




## 4.1 Estructura de la IP<sub>5</sub> 2-K de *A. thaliana*. Complejos con sustratos (IP<sub>5</sub> e IP<sub>5</sub>/AMPPNP) y productos (IP<sub>6</sub> e IP<sub>6</sub>/ADP)

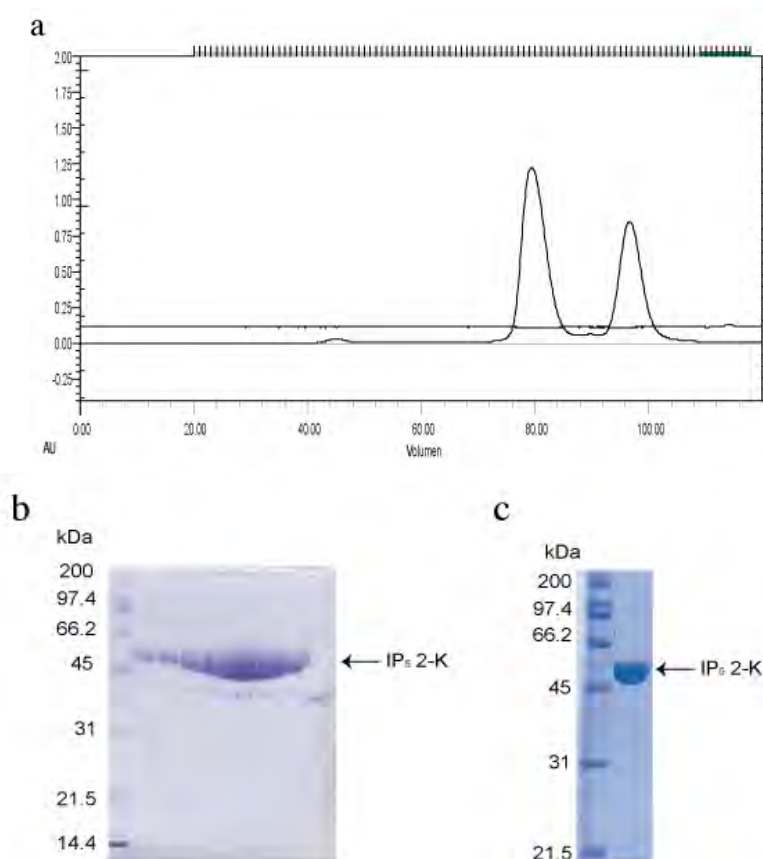
### 4.1.1 Obtención de muestra para estudios de cristalización

La purificación de IP<sub>5</sub> 2-K se realizó partiendo de cuatro litros de cultivo. El primer paso de purificación (columna de *Sefarosa* CL-6) permitió aislar con éxito la proteína de fusión LSL-IP<sub>5</sub> 2-K, cuyo peso molecular corresponde al de la suma de ambas (**Figura 4.1a**). Posteriormente, la columna de heparina incrementó significativamente el grado de pureza de la proteína de fusión (**Figura 4.1b**). La proteína eluye en el 61% del tampón A, correspondiente a 400 mM de NaCl. El corte con TEV proteasa en estas condiciones es prácticamente total (**Figura 4.1c**).



**Figura 4.1. Purificación de IP<sub>5</sub> 2-K.** a) SDS-PAGE mostrando la proteína obtenida tras el paso de captura con *Sefarosa* CL-6. Se observan las fracciones de proteína obtenida tras la elución con 200 mM de lactosa. b) Proteína obtenida tras la columna de heparina. Fracciones correspondientes a la elución de la columna de heparina con un gradiente de NaCl entre 0-1 M. c) Corte con proteasa TEV. Los carriles muestran, 1: muestra sin incubar con proteasa TEV. 2: muestra tras una incubación con proteasa TEV durante 4 horas. 3: muestra tras una incubación con proteasa TEV 12 horas. Como se aprecia en esta última fracción el corte es prácticamente total.

En la cromatografía de exclusión molecular se obtuvieron dos picos, uno correspondiente a IP<sub>5</sub> 2-K (elución en el mL 77) y un segundo correspondiente a LSL (elución en el mL 95) (**Figura 4.2a y b**). En el último paso de purificación (columna de *Sefarosa* CL-6) se eliminaron los posibles restos de proteína de fusión. Se obtuvieron un total de 40 mg de proteína, lo que correspondería a un rendimiento final de 10 mg/L de cultivo (**Figura 4.2c**). La muestra se concentró a 10 mg/mL para los experimentos de cristalización.



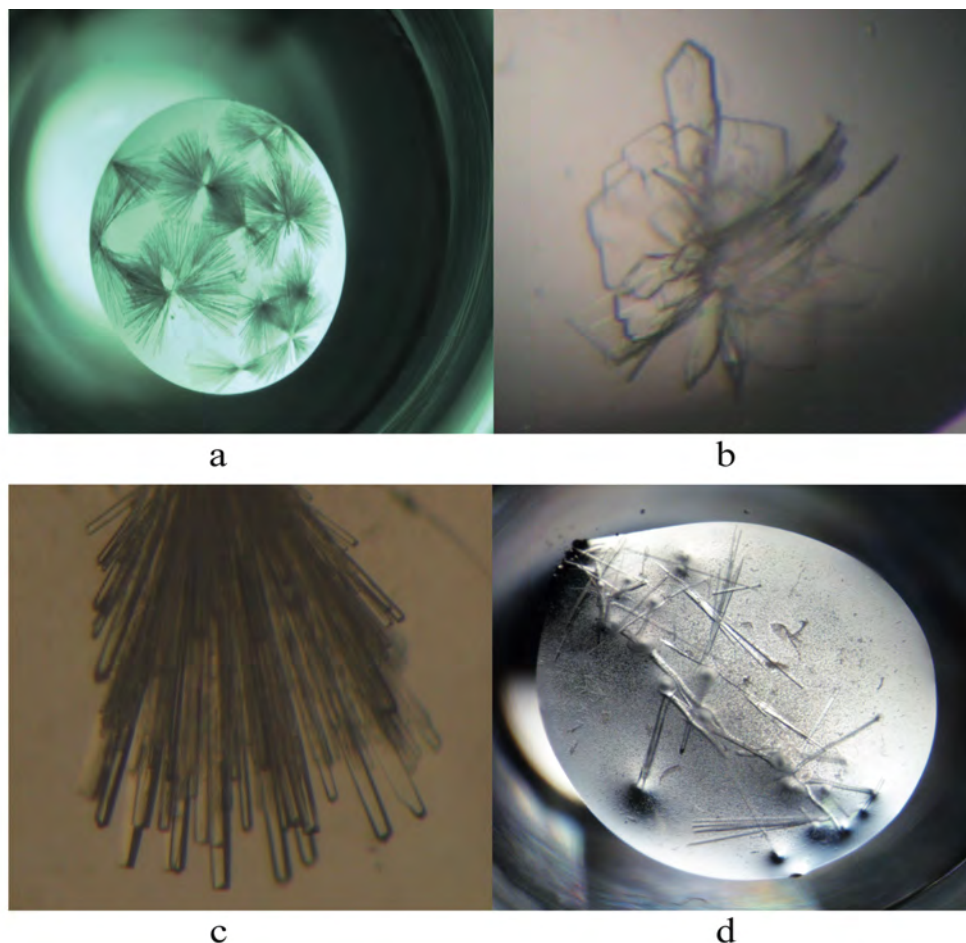
**Figura 4.2. Pasos finales de la purificación de IP<sub>5</sub> 2-K.** a) Perfil de la cromatografía de exclusión molecular de IP<sub>5</sub> 2-K. b) SDS PAGE de la proteína obtenida tras la exclusión molecular. Las fracciones mostradas corresponden al primer pico del cromatograma mostrado en a). c) SDS PAGE de la muestra final durante el proceso de concentración, con la que se llevó a cabo la cristalización de la enzima.

### 4.1.2 Cristalización de la IP<sub>5</sub> 2-K y complejos con inositido

Los primeros ensayos de cristalización de la muestra se realizaron sin incubar ésta con ningún sustrato o producto. Pese al gran número de condiciones ensayadas, no se lograron obtener cristales.

A continuación, y previamente a los ensayos de cristalización, se incubó la muestra con el producto de reacción IP<sub>6</sub> (2 mM). IP<sub>5</sub> 2-K en complejo con IP<sub>6</sub> cristalizó a pHs distintos: a pH 6.5 (20% (m/v) PEG 3350, 0.1 M bis-tris propano pH 6.5) se obtenían conjuntos de placas (**Figura 4.3b**) y a pH 5.9 (22% (m/v) PEG 3350, 0.1 M Bis-Tris pH 5.9) se obtenían conjuntos de agujas (**Figura 4.3a**). Los cristales en forma de placas no permitieron el indexado de su patrón de difracción decidiéndose por tanto continuar optimizando los cristales en forma de agujas.

Los mejores cristales se obtuvieron mediante la inclusión de ambos productos de reacción en la solución (2 mM  $\text{IP}_6$  y 2 mM ADP) (**Figura 4.3c**). Se empleó también la técnica de sembrado, lográndose separar los conjuntos de agujas, si bien los cristales obtenidos eran de menor grosor que los obtenidos sin el empleo de esta técnica (**Figura 4.3d**), por lo que se descartó su uso.



**Figura 4.3. Cristalización de  $\text{IP}_5$  2-K.** a) Agujas obtenidas con 22% (m/v) PEG 3350, 0.1 M Bis-Tris pH 5.9, 2 mM  $\text{IP}_6$ . b) Placas obtenidas con 20% (m/v) PEG 3350, 0.1 M bis-tris propano pH 6.5. c) Agujas obtenidas con 21% (m/v) PEG 3350, 0.1 M Bis-Tris pH 5.9, 2 mM  $\text{IP}_6$  y 2 mM ADP. d) Agujas obtenidas con 18% (m/v) PEG 3350, 0.1 M Tis-Tris pH 5.9, 2 mM  $\text{IP}_6$  y 2 mM ADP, empleando la técnica de sembrado.

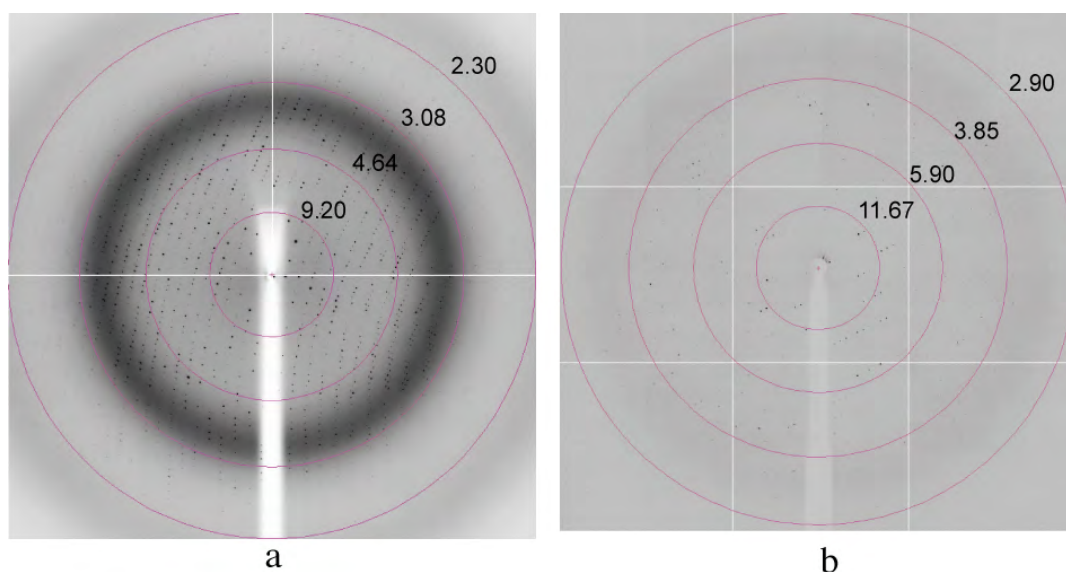
A partir de estas condiciones de cristalización y tras el ajuste de condiciones para cada uno de ellos, se cristalizaron los complejos obtenidos con sustratos y análogos de los mismos, ( $\text{IP}_5$  2-K/ $\text{IP}_5$ ,  $\text{IP}_5$  2-K/ $\text{IP}_5$ /AMPPNP,  $\text{IP}_5$  2-K/I(3,4,5,6) $\text{P}_4$ /AMPPNP). Las condiciones finales de cristalización de cada complejo se muestran en la (**Tabla 4.1**). En todos los casos los cristales aparecían en 5 días, y continuaban creciendo durante 2 semanas, hasta alcanzar una dimensión promedio de 0.3 x 0.05 x 0.03 mm.

**Tabla 4.1. Condiciones de cristalización de los diferentes complejos de IP<sub>5</sub> 2-K.**

Complejo	Condición
IP <sub>5</sub> 2-K/IP <sub>5</sub>	23% PEG 3350, 0.1 M Bis-Tris pH 5.9
IP <sub>5</sub> 2-K/IP <sub>5</sub> /AMPNP	22% PEG 3350, 0.1 M Bis-Tris pH 5.9
IP <sub>5</sub> 2-K/IP <sub>6</sub> /ADP	21% PEG 3350, 0.1 M Bis-Tris pH 5.9
IP <sub>5</sub> 2-K/IP <sub>6</sub>	22% PEG 3350, 0.1 M Bis-Tris pH 5.9
IP <sub>5</sub> 2-K/IP <sub>6</sub> /ADP (1 molécula/u.a)	21% PEG 3350, 0.1 M Bis-Tris pH 5.9
IP <sub>5</sub> 2-K/I(3,4,5,6)P <sub>4</sub> /AMPNP (1 molécula/u.a)	22% PEG 3350 y 0.1 M Bis-Tris pH 5.9

### 4.1.3 Resolución estructural de la IP<sub>5</sub> 2-K y construcción de modelos.

Como figura en la sección de Materiales y Métodos, se obtuvieron conjuntos de datos de los diferentes complejos cristalizados (**Figura 4.4a**). Los cristales derivados de átomo pesado que permitieron la resolución estructural fueron aquellos que habían sido sumergidos en acetato de plomo 100 mM durante 25 minutos (**Figura 4.4b**).



**Figura 4.4. Patrones de difracción de IP<sub>5</sub> 2-K.** a) Patrón obtenido a partir del cristal con IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP. b) Patrón de difracción obtenido a partir del cristal derivado con plomo IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP/Pb<sup>2+</sup>.

En la **Tabla 4.2** pueden observarse los parámetros del procesamiento correspondiente a los diferentes conjuntos de datos nativos. Como puede observarse, el cambio de agente crioprotector (de glicerol a un aumento en la concentración de PEG) propició una disminución del tamaño de la celda unidad, siendo notable la reducción de 11 Å del

eje *c*.

Se obtuvieron dos tipos de cristales distintos, con grupo espacial  $P2_12_12_1$  o  $P2_12_12$ . Los cristales pertenecientes al grupo  $P2_12_12$  presentaron un menor tamaño de la celdilla unidad y una molécula en la unidad asimétrica en lugar de dos. Esta variación, dependía de la preparación de proteína usada, probablemente debida al contenido en  $Zn^{2+}$  de la muestra.

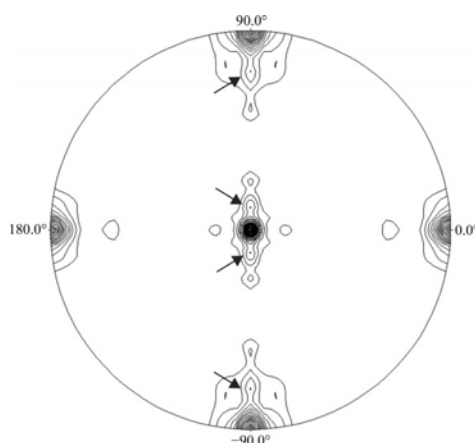
Los parámetros del procesamiento de datos del derivado con plomo y el faseado de la estructura se encuentran resumidos en la **Tabla 4.3**. La calidad de los datos de este derivado es menor que la de los datos correspondientes al mismo tipo de cristal nativo (3.2 Å frente a 2.2 Å). Como puede observarse hay una ligera reducción del tamaño de la celdilla unidad, con lo que el contenido de solvente de estos cristales es menor que el de los cristales nativos (43.2% considerando dos moléculas en la unidad asimétrica).

**Tabla 4.2. Procesado de datos de los diferentes complejos de IP<sub>5</sub> 2-K.** Los valores entre paréntesis corresponden a la última capa de resolución.

	IP <sub>6</sub>	IP <sub>6</sub>	IP <sub>6</sub> /ADP	IP <sub>5</sub> /AMPPNP	IP <sub>5</sub>	I(3,4,5,6)P <sub>4</sub> /AMPPNP	IP <sub>6</sub> /ADP (1 mol./u.a.)
Grupo espacial	$P2_12_12_1$					$P2_12_12$	
Moléculas por u.a.	2					1	
Celdilla (Å)	57.66 117.52 150.44	57.66 112.50 139.72	58.12 113.59 142.48	57.42 113.43 140.47	58.76 112.74 144.45	61.62 65.08 122.49	62.11 65.66 121.79
Crioprotección	20% Glicerol	35% PEG 3350					
Línea de luz	ID14-2 (ESRF)	ID14-1 (ESRF)	ID14-1 (ESRF)	ID14-4 (ESRF)	ID14-4 (ESRF)	ID14-1 (ESRF)	ID14-4 (ESRF)
Longitud de onda (Å)	0.9330	0.9343	0.9334	0.9395	0.9538	0.9334	0.9794
Resolución (Å)	74.53-3.20	112.50-3.10	71.24-2.20	88.25-2.20	88.74-2.90	57.47-2.93	60.9-2.40
R <sub>merge</sub> (%)	15.6 (47.1)	11.6 (43.6)	13.5 (53.2)	9.9 (46.3)	13.0 (44.6)	14.6 (39.4)	13.7 (50.7)
Reflexiones únicas	18140	17187	48686	47247	22203	11132	19340
Redundancia (%)	5.7 (5.9)	7.6 (7.9)	7.1 (5.7)	7.6 (7.6)	6.9 (7.0)	7.0 (6.9)	5.5 (5.6)
Compleitud (%)	100 (100)	100 (100)	99.8 (99.8)	99.6 (99.6)	100 (100)	100 (100)	96.4 (91.0)
$\langle I/\sigma(I) \rangle$	12.7 (2.9)	13.7 (4.9)	10.4 (3.0)	16.3 (4.4)	13.7 (3.8)	13.4 (4.9)	9.6 (3.1)
Wilson B (Å <sup>2</sup> )	80.7	54.7	29.0	34.6	41.1	39.2	31.9

$$R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$$

En el caso de los cristales que presentaban dos moléculas en la unidad asimétrica, se analizaron los picos de la función de autorotación, lo que permitió observar la presencia de un eje de simetría no cristalográfico a lo largo del plano formado por los ejes *bc*. Una de las secciones de la proyección estereográfica ( $\kappa=180^\circ$ ) puede observarse en la **Figura 4.5**.



**Figura 4.5. Representación de la función de autorrotación de los cristales de IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP (2 moléculas/u.a.).** Se emplearon datos entre 2.2 y 15 Å y un radio de integración de 25 Å. Se presenta la sección  $\kappa=180^\circ$  proyectada a lo largo del eje  $c$ , donde  $\varphi = 0^\circ$  y  $\varphi = 90^\circ$  corresponde a los ejes  $a$  y  $b$  respectivamente. Las flechas indican un eje de simetría no cristalográfica a través del plano  $bc$ .

**Tabla 4.3. Procesado de datos del derivado con plomo de IP<sub>5</sub> 2-K.** Los valores entre paréntesis corresponden a la última capa de resolución.

	Pico	Inflexión	Remoto
<b>Recogida de datos</b>			
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>		
Moléculas por u.a.	2		
Celdilla (Å)	58.06 110.97 138.69		
Línea de luz	ID23-1 (ESRF)		
Longitud de onda (Å)	0.9478	0.9503	0.9442
Resolución (Å)	138.67 - 3.20	140.02 - 3.20	138.70 - 3.20
R <sub>merge</sub> (%)	16.8 (52.0)	19.3 (66.7)	18.0 (57.3)
Reflexiones únicas	16948	17169	16898
Redundancia (%)	6.7 (6.7)	6.7 (6.9)	6.7 (6.8)
Redundancia anómala (%)	3.6 (3.5)	3.6 (3.5)	3.6 (3.5)
Compleitud (%)	100.0 (100.0)	100.0 (100.0)	99.9 (99.6)
Compleitud anómala (%)	99.7 (98.9)	99.9 (99.7)	99.9 (99.5)
$\langle I/\sigma(I) \rangle$	12.1 (3.8)	11.0 (3.0)	12.1 (3.7)
Wilson B (Å <sup>2</sup> )	53.8	56.6	53.7
<b>Determinación de fases</b>			
CC <sub>all</sub> /CC <sub>weak</sub> /PATFOM	42.41/19.67/33.65	-	-
Resolución (Å)	86.65 - 3.2	-	-
Dispersores anómalos	4		
Figura de Mérito (FOM) (Acéntricas/Céntricas/Totales)	0.4271/0.2441/0.4012	-	-
<b>R<sub>Cullis</sub></b>			
Anómalo (acéntricas)	0.884	0.938	0.889
Isomorfo (acéntricas/céntricas/totales)	-	0.769/0.794/0.772	0.748/0.797/0.754
<b>Poder de faseado</b>			
Anómalo (acéntricas)	0.856	0.568	0.843
Isomorfo (acéntricas/céntricas/totales)	-	0.421/0.370/0.414	0.019/0.016/0.018

$$R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$$

Como se explicó en la sección de Materiales y Métodos, las estructuras de los complejos que presentaban isomorfismo entre sí y que no estaban derivatizados con



átomos pesados (esto es IP<sub>5</sub> 2-K/IP<sub>5</sub>, IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP e IP<sub>5</sub> 2-K/IP<sub>6</sub>) fueron obtenidas mediante un ajuste de cuerpo rígido del modelo obtenido experimentalmente y el posterior cálculo de mapas de diferencia.

En el caso de los complejos que sólo tenían una molécula en la unidad asimétrica (IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP e IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP), éstos fueron resueltos mediante Reemplazamiento Molecular.

Los valores geométricos y estereoquímicos obtenidos tras el refinamiento se encontraban dentro de los valores esperables (Tabla 4.4).

**Tabla 4.4.** Estadísticas del refinamiento de los diferentes conjuntos de datos de IP<sub>5</sub> 2-K.

	IP <sub>6</sub> /ADP/Pb <sup>2+</sup>	IP <sub>5</sub> /AMPPNP	IP <sub>6</sub> /ADP	IP <sub>6</sub>	IP <sub>5</sub>	I(3,4,5,6)P <sub>4</sub> /AMPPNP	IP <sub>6</sub> /ADP (1 mol./u.a.)
<b>Refinamiento</b>							
Resolución (Å)	86.71-3.20	140.03-2.20	88.73-2.20	112.50-3.10	88.74-2.90	57.47-2.93	60.90-2.40
R <sub>work</sub> (%)	21.0	22.4	22.8	24.1	22.4	22.0	24.3
R <sub>free</sub> (%)	26.8	27.7	26.9	29.8	29.0	28.0	29.2
<b>Número de átomos (no H)</b>							
Totales	6851	6991	7015	6751	6668	3320	3422
Proteína	6719	6867	6694	6671	6602	3260	3297
Ligando	124	124	124	72	64	59	62
Iones	6	4	6	2	2	2	2
Moléculas de agua	0	259	193	0	0	0	60
<b>Desviaciones cuadráticas medias (rmsd)</b>							
Enlaces (Å)	0.011	0.006	0.007	0.008	0.010	0.005	0.006
Ángulos (°)	1.462	1.109	1.149	1.190	1.371	1.060	1.133
<b>Factor térmico B(Å<sup>2</sup>)</b>							
Total	32.8	29.0	34.6	54.7	41.1	29.1	27.4
Proteína	32.4	29.3	35.2	55.8	41.3	29.05	27.6
Ligandos	13.2	18.2	21.3	29.1	21.7	18.34	24.3
<b>Diagrama de Ramachandran</b>							
Residuos en la zona más favorable (%)	86.1	91.2	91.0	89.9	91.4	92.4	90.2
<b>Residuos no modelados</b>							
Cadena A	1:2	1:2	1:2	1:2	1:2	47:58	48:58
	50:58	46:58	49:58	47:58	50:58	155:158	156:158
	336:341	156:157	156:157	157	154:159	378:386	379:387
	438:451	336:341	336:340	336:341	252:253	437:451	439:451
Cadena B	1:2	1:2	1:2	1:2	1:2	-	-
	54:56	46:56	54:56	47:56	53:57	-	-
	153:160	156:158	156:158	381:385	154:160	-	-
	381:382	380:383	380:383	442:451	438-451	-	-
Cadena C	439:451	439:451	439:451	439:451	439:451	-	-
	439:451	439:451	439:451	439:451	439:451	-	-
	439:451	439:451	439:451	439:451	439:451	-	-
	439:451	439:451	439:451	439:451	439:451	-	-

$$R_{\text{work}}/R_{\text{free}} = \sum_{hkl} |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{calc}}| / \sum_{hkl} F_{hkl}^{\text{obs}}$$



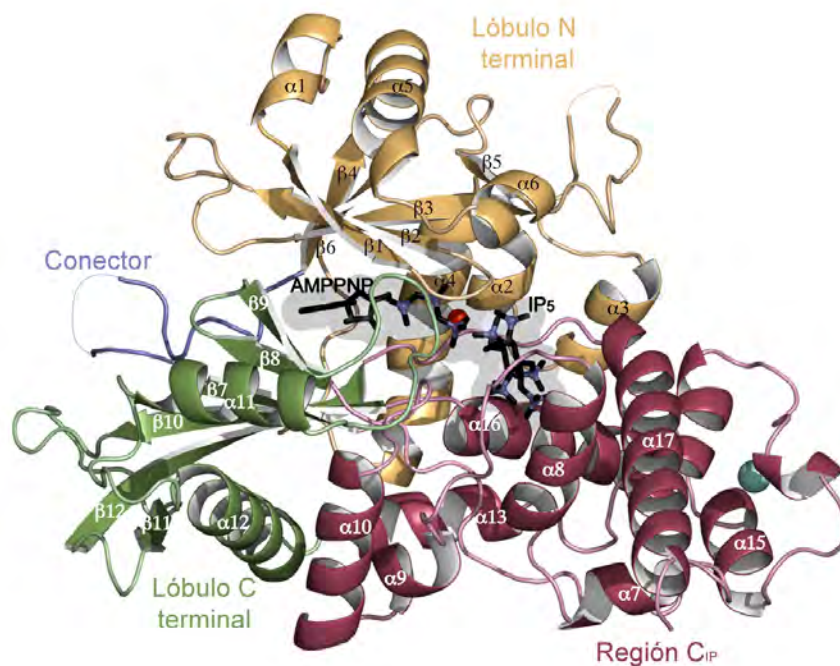
En todos los casos, se obtuvieron mapas de buena calidad que permitieron asignar los ligandos correspondientes (**Figura 4.6**).



**Figura 4.6.** Mapas de densidad electrónica de los cristales correspondientes a los complejos de IP<sub>5</sub> 2-K. a) Mapa de densidad electrónica a 1σ de IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP. b) Densidad electrónica a 1σ alrededor del IP<sub>5</sub> en el complejo IP<sub>5</sub> 2-K/IP<sub>5</sub>. c) Densidad electrónica a 1σ alrededor del IP<sub>6</sub> en el complejo IP<sub>5</sub> 2-K/IP<sub>6</sub>. d) Densidad electrónica a 1σ alrededor del IP<sub>5</sub> y AMPPNP en el complejo IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP. e) Densidad electrónica a 1σ alrededor del IP<sub>6</sub> y ADP en el complejo IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP. f) Densidad electrónica a 1σ alrededor del I(3,4,5,6)P<sub>4</sub> y AMPPNP en el complejo IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP. g) Densidad electrónica a 5σ en torno al átomo de Zn<sup>2+</sup> encontrado en IP<sub>5</sub> 2-K.

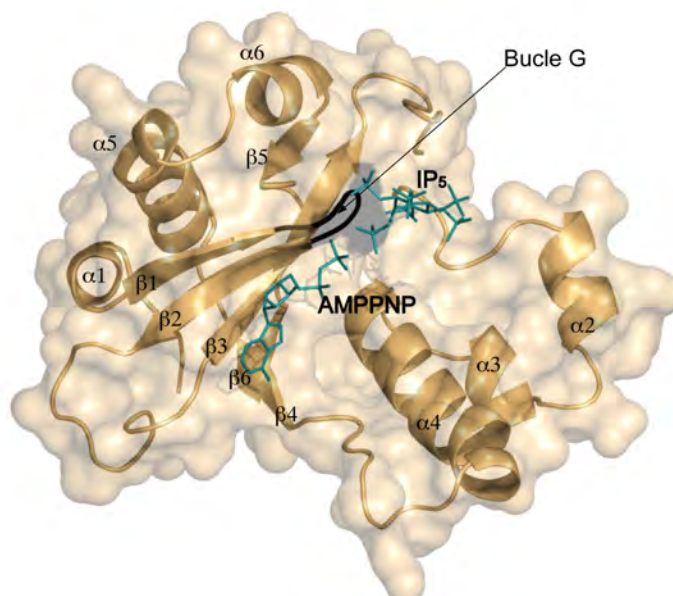
#### 4.1.4 Estructura de la IP<sub>5</sub> 2-K: descripción general del plegamiento

IP<sub>5</sub> 2-K presenta dos lóbulos: un lóbulo N-terminal (residuos 1-152; lóbulo N<sub>t</sub>) y lóbulo C-terminal (residuos 161-432; lóbulo C<sub>t</sub>) los cuales se encuentran unidos por un bucle conector (153-160) (**Figura 4.7**).



**Figura 4.7. Organización general de la IP<sub>5</sub> 2-K de *A. thaliana*.** IP<sub>5</sub> 2-K se muestra mediante una representación de cintas, en la que el lóbulo N<sub>t</sub> se muestra en color naranja, el lóbulo C<sub>t</sub> en verde y en magenta (región C<sub>IP</sub>) y en violeta el conector entre ambos lóbulos. Se muestran también los ligandos (AMPPNP e IP<sub>5</sub>) en color negro, resaltando los fosfatos en azul, el ión Mg<sup>2+</sup> mediante una esfera roja y el ión Zn<sup>2+</sup> mediante una esfera azul claro.

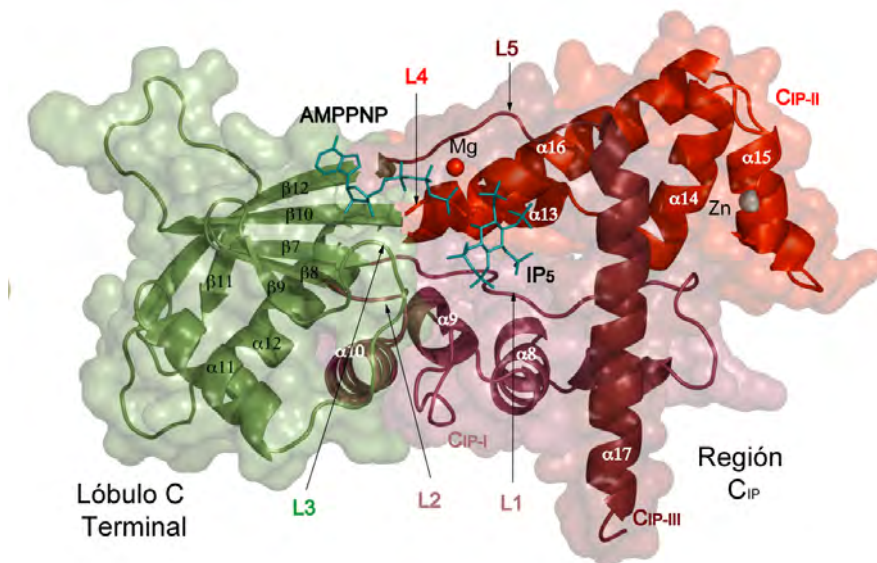
El lóbulo N<sub>t</sub> (**Figura 4.7** y **Figura 4.8**) presenta un plegamiento  $\alpha+\beta$ . El núcleo consiste en una lámina  $\beta$  antiparalela compuesta de cinco hebras  $\beta$  en orden 1-2-3-6-4. Las hebras  $\beta 3-\beta 4$  se encuentran conectadas por un segmento de carácter  $\alpha$ -helicoidal formado por un bucle y tres hélices dispuestas de forma antiparalela ( $\alpha 2-\alpha 4$ ). Las hebras  $\beta 4-\beta 6$  se encuentran conectadas por una hélice  $\alpha$  ( $\alpha 5$ ) que cruza por completo la lámina  $\beta$ , seguida de una pequeña hélice antiparalela ( $\alpha 6$ ) y una pequeña hebra  $\beta$  ( $\beta 5$ ). La  $\beta 5$  interacciona de forma antiparalela con la porción C-terminal de  $\beta 3$ .



**Figura 4.8. Organización del lóbulo N-terminal ( $N_t$ ) de  $IP_5$  2-K.** El lóbulo  $N_t$  se muestra mediante una combinación de cintas y superficie, empleando el mismo color que en la **Figura 4.7**. Los ligandos se destacan en color cian y el bucle G está señalado en color negro.

El lóbulo C-terminal (lóbulo  $C_t$ ) posee también un plegamiento  $\alpha+\beta$ , pero presenta un tamaño mayor (prácticamente el doble) que el  $N_t$  (**Figura 4.7** y **Figura 4.9**). Las diferentes hebras  $\beta$  se encuentran conectadas en orden 9-8-7-10-12-11 formando una lámina  $\beta$  antiparalela. La mayor parte de las hélices definen la región  $C_{IP}$ , denominada así porque constituye el sitio de unión al IP o inositido. Esta región  $C_{IP}$  se encuentra a su vez formada por tres subregiones no contiguas en secuencia. La primera de ellas,  $C_{IP-I}$  (residuos 169-239) conecta las hebras  $\beta 7$  y  $\beta 8$  y está formada por tres hélices  $\alpha$  ( $\alpha 7$ - $\alpha 10$ ). La segunda región,  $C_{IP-II}$  (residuos 250-368), conecta la  $\alpha 12$  con la  $\beta 10$ , y está formada por otras cuatro hélices  $\alpha$  ( $\alpha 13$ - $\alpha 16$ ). La tercera región,  $C_{IP-III}$  (residuos 407-439) está constituida por una única hélice de gran tamaño,  $\alpha 17$ , la cual actúa como soporte para las regiones  $C_{IP-I}$  y  $C_{IP-II}$ .

La región  $C_{IP}$  y el resto del lóbulo C-terminal están conectados mediante diferentes bucles: L1 (residuos 167-182), L2 (residuos 235-240), L4 (residuos 368-370) y L5 (residuos 406-413). Estos bucles (L1, L2, L4 y L5) albergan residuos clave para la enzima, como se verá más adelante. El bucle L3 (residuos 250-262) conecta elementos del lóbulo  $C_t$  no pertenecientes a la región  $C_{IP}$ .



**Figura 4.9. Organización del lóbulo C-terminal ( $C_t$ ) de  $IP_5$  2-K.** El lóbulo  $C_t$  se muestra mediante una combinación de cintas y superficie, empleando el mismo código de colores que en la **Figura 4.7**. Cada zona de la región  $C_{IP}$  se señala con un tono de rojo diferente. Los ligandos se destacan en color azul. Los iones  $Mg^{2+}$  y  $Zn^{2+}$  se muestran como esferas roja y gris respectivamente.

Como se ha mencionado, en la mayoría de los complejos obtenidos los cristales presentaban dos moléculas en la unidad asimétrica, existiendo una relación de pseudosimetría binaria entre ambas moléculas. Existen pocos contactos entre los dos monómeros (un total de 16 enlaces de hidrógeno), siendo la superficie enterrada total de  $1101 \text{ \AA}^2$ . Esto sugiere que el dímero observado en el cristal no se encuentra presente en solución, lo que es compatible con los resultados obtenidos mediante la cromatografía de exclusión molecular empleada para purificar la enzima. No se observaron cambios entre los centros activos de las dos moléculas presentes en la unidad asimétrica, presentando la superposición de ambas moléculas una desviación cuadrática media de  $0.704 \text{ \AA}$ .

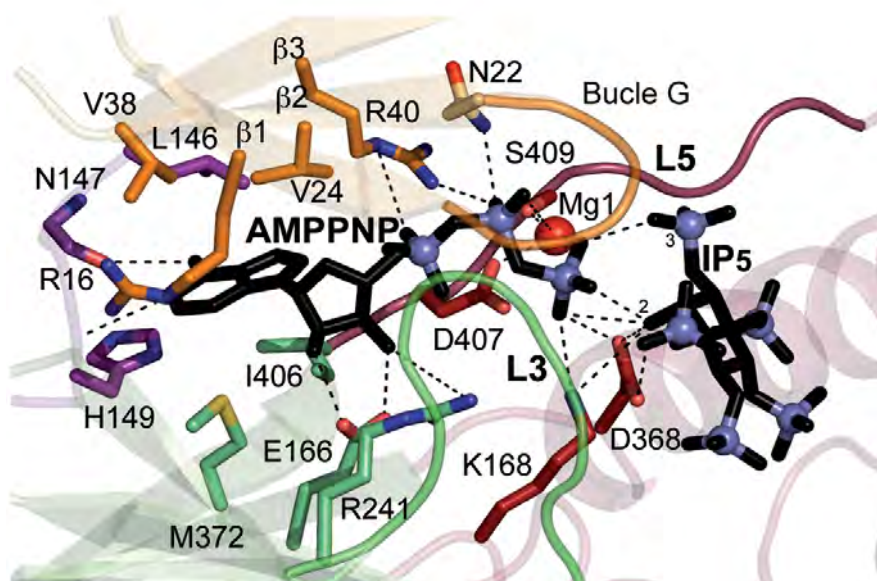
No se observaron cambios en el plegamiento general de la enzima entre los diferentes complejos cristalizados, no mostrando éstos diferencias estructurales significativas (rmsd máxima de  $0.479 \text{ \AA}$ ).



## 4.1.5 El centro activo de la IP<sub>5</sub> 2-K

### 4.1.5.1 Sitio de unión al nucleótido

El nucleótido se encuentra presente en las estructuras de los complejos ternarios obtenidos, bien como análogo no hidrolizable del sustrato ATP (complejos IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP e IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP) o producto de reacción (complejos IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP) (**Figura 4.6**).



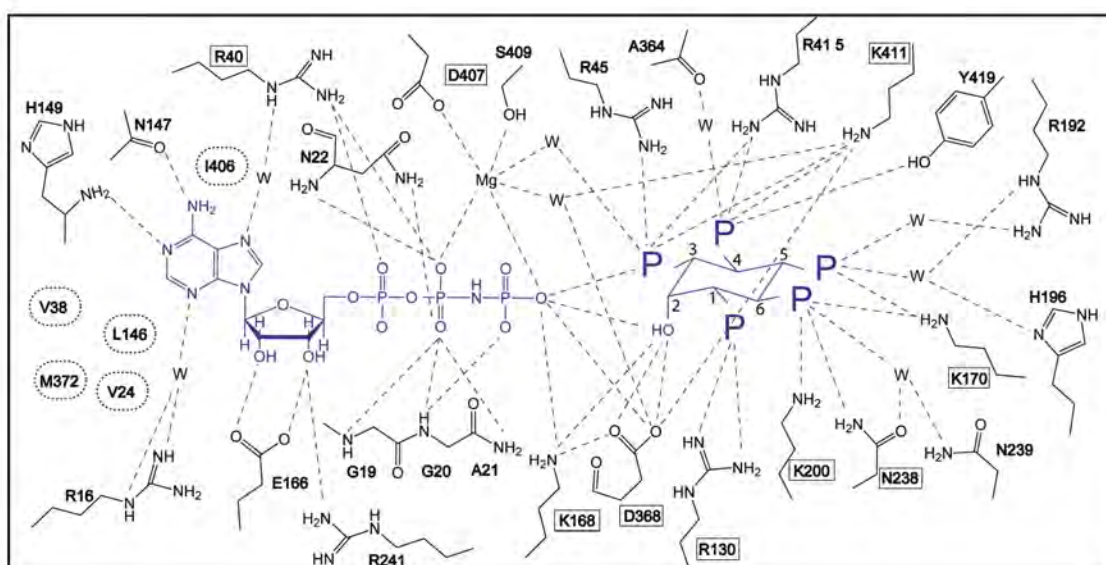
**Figura 4.10. Interacciones del nucleótido en el centro activo de IP<sub>5</sub> 2-K.** Detalle del centro activo siguiendo el código de colores de la **Figura 4.7**. La proteína se representa en un modelo de cintas transparente, resaltándose los residuos implicados en la unión de AMPPNP en el mismo color que la región de donde proceden, con los átomos de oxígeno, nitrógeno y azufre en color rojo, azul y amarillo respectivamente. Los ligandos AMPPNP e IP<sub>5</sub> se representan en color negro con los fosfatos como esferas de color azul. Las interacciones formadas se representan con líneas de puntos discontinuas.

El nucleótido (AMPPNP) está unido por los lóbulos N<sub>t</sub> y C<sub>t</sub>, en la interfase creada entre las láminas β de ambos lóbulos (**Figura 4.7** y **Figura 4.10**). Un bolsillo hidrofóbico formado con residuos de ambos lóbulos (Arg16, Val24, Val38, Leu146, His149, Met372 e Ile406) aloja la adenina y la ribosa. También se observan múltiples interacciones polares; por un lado el anillo de adenina coordina con dos residuos del bucle conector, Asn147 e His149, mediante puentes de hidrógeno, y por otro lado, los átomos de oxígeno de la ribosa coordinan con Glu166 y Arg241. En cuanto a los fosfatos del nucleótido, Pα es coordinado principalmente por Arg40. Pβ está coordinado por residuos provenientes de un bucle del lóbulo N<sub>t</sub> (denominado bucle G o bucle rico en glicinas) el cual se localiza entre las hebras β1 y β2. En concreto, el bucle

G coordina el P $\beta$  mediante las cadenas principales de Gly19, Gly20, Ala21 y Asn22. Este fosfato está también coordinado por Arg40. Por último, el P $\gamma$  coordina con Gly20 del anterior bucle, Lys168 (L1), Asp368 (L4) y con Asp407 y Ser409 (L5) mediante un ión Mg<sup>2+</sup>, el cual a su vez también coordina con el P $\beta$  del nucleótido. Todas las interacciones descritas se encuentran también representadas esquemáticamente en la **Figura 4.11**.

Cabe destacar en este punto que el bucle L3 junto con el bucle G parecen cerrar el centro activo cubriendo el nucleótido y ayudando además a fijar los fosfatos en el centro activo (**Figura 4.10**).

Las interacciones descritas se encuentran representadas en la **Figura 4.11**. En resumen los residuos encargados de coordinar el nucleótido provienen tanto del lóbulo N<sub>t</sub> ( $\beta$ 1,  $\beta$ 2,  $\beta$ 3 y bucle G) como del conector entre los lóbulos N<sub>t</sub> y C<sub>t</sub> y del propio C<sub>t</sub> (bucles L1, L4 y L5), encargándose el bucle L3 y el bucle G de cerrar el sitio de unión al nucleótido.



**Figura 4.11. Diagrama de interacciones entre IP<sub>5</sub> 2-K y sustratos.** En color negro se representan los residuos de IP<sub>5</sub> 2-K y en color azul los ligandos AMPPNP e IP<sub>5</sub>. Las interacciones intermoleculares se representan mediante líneas de puntos discontinuas.

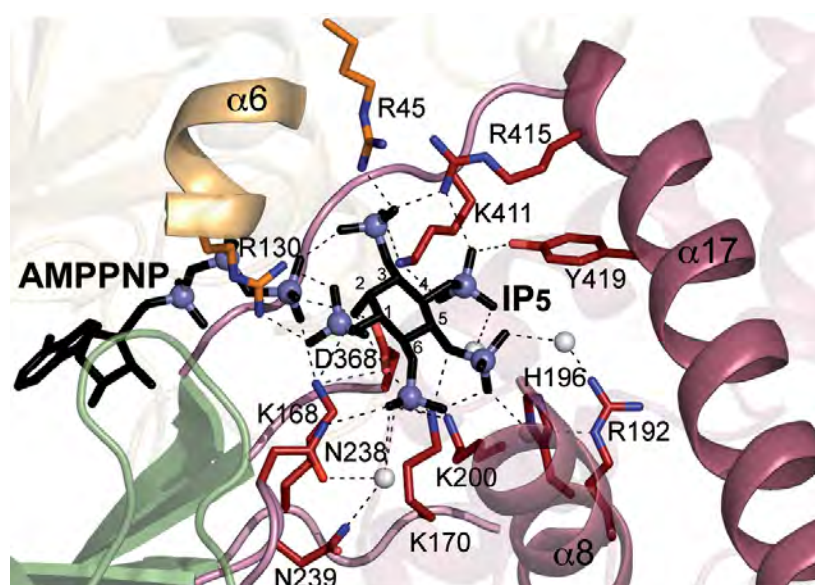
Todas las interacciones observadas en el complejo IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP también se producen en el complejo ternario IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP.

En cuanto al ADP, las interacciones que mantiene con el sitio de unión al nucleótido son prácticamente las mismas que las descritas para el AMPPNP, salvo por aquellas que

varían debido a la transferencia del Py durante la catálisis, las cuales serán explicadas más adelante (Sección 4.1.5.3.).

### 4.1.5.2 Sitio de unión al inositido

El inositido está presente en todas las estructuras obtenidas, bien como sustrato (complejos  $IP_5$  2-K/ $IP_5$ /AMPPNP,  $IP_5$  2-K/ $IP_5$  e  $IP_5$  2-K/ $I(3,4,5,6)P_4$ /AMPPNP) o producto (complejo  $IP_5$  2-K/ $IP_6$ /ADP e  $IP_5$  2-K/ $IP_6$ ) (**Figura 4.6**). Tanto el sustrato ( $IP_5$ ) como el producto ( $IP_6$ ) muestran el mismo patrón de interacciones con  $IP_5$  2-K, encontrándose unidos a la enzima mediante todos sus grupos fosfato (**Figura 4.12**).

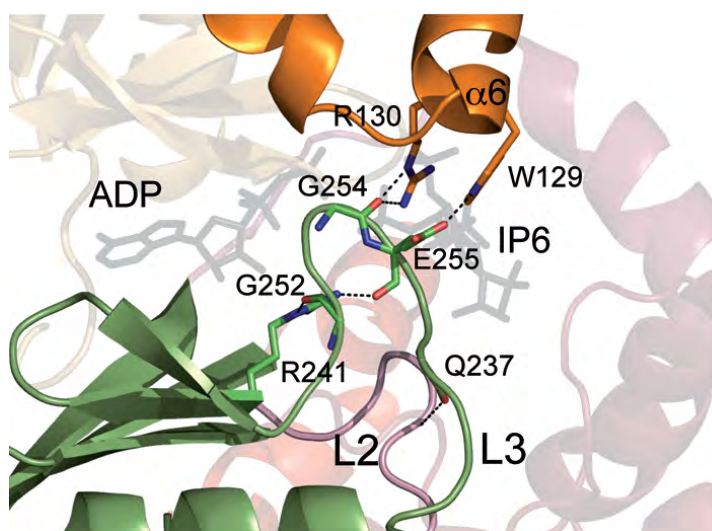


**Figura 4.12.** Interacciones del inositido en el centro activo de  $IP_5$  2-K. La estructura de  $IP_5$  2-K se representa mediante un esquema de cintas. El color de la proteína, residuos y ligandos sigue el código de la **Figura 4.10**.

En cuanto a  $IP_5$ , el fosfato 1 (P1) está coordinado por la Arg130 ( $\alpha 6$ ), siendo el único fosfato que coordina únicamente con el lóbulo  $N_t$ . El P3 está unido tanto al lóbulo  $N_t$  (Arg45) como a la región  $C_{IP}$  (Lys411 y Arg415). El resto de grupos fosfato están exclusivamente coordinados por residuos del lóbulo C, y en particular de la región  $C_{IP}$ . Así el P4 está coordinado por Lys411, Arg415 y Tyr419, el P5 por Lys170 y a través de moléculas de agua con Arg192 e His196, y el P6 por Lys170, Lys200 y Asn238. Finalmente, el grupo hidroxilo en la posición 2 del anillo interacciona con la Lys168 y el Asp368 del lóbulo  $C_t$ . En el caso del  $IP_6$ , el P2 produce las mismas interacciones con la enzima que el Py del ATP (Sección 4.2.5.3). Muchos de los residuos descritos se localizan en los bucles L1, L2, L4 y L5.

Las interacciones explicadas se conservan en los diferentes complejos cristalizados, incluyendo el complejo con  $I(3,4,5,6)P_4/AMPPNP$ . El centro activo de este complejo se discutirá más adelante.

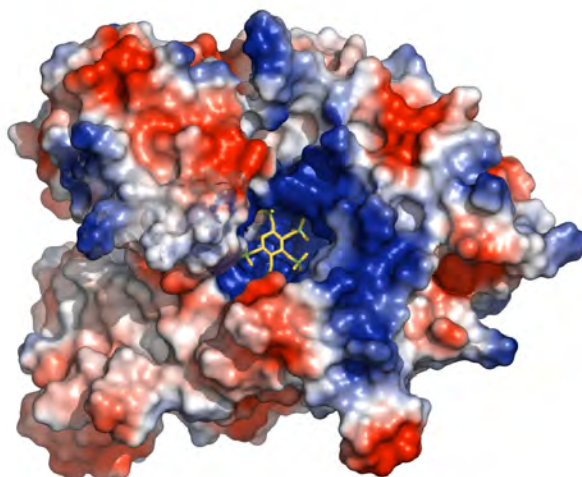
Si bien el bucle L3 no une directamente el inositol, sí interacciona con elementos de reconocimiento del IP. En concreto, L3 interacciona por un lado con la hélice  $\alpha_6$  del lóbulo N, en concreto formando los puentes Gly254-Arg130 y Glu255-Trp129, y por otro lado con el bucle L2 del lóbulo C (Gly252-Arg241 y Glu255-Arg241) (**Figura 4.13**). Estos elementos,  $\alpha_6$  y L2, interaccionan con P1 y P6 respectivamente. Por tanto, el bucle L3 parece estar también implicado en la estabilización de la unión a inositido así como en la interacción entre los lóbulos  $N_t$  y  $C_t$ .



**Figura 4.13. Interacción del bucle L3 con elementos de reconocimiento del IP.** La estructura de  $IP_5$  2-K se muestra mediante un esquema de cintas, destacándose los residuos implicados en la interacción del bucle L3 con las regiones  $\alpha_6$  y L2. Los enlaces formados se representan con líneas de puntos discontinuas. El color de los diferentes residuos está de acuerdo a lo explicado en la **Figura 4.7**. El bucle L2 se destaca en color rosa.

En resumen, la mayoría de los residuos que coordinan los fosfatos provienen de la región  $C_{IP}$ , aunque algunos son aportados por el lóbulo  $N_t$ . Todos estos residuos dotan al sitio de unión al sustrato de  $IP_5$  2-K de una fuerte carga positiva, necesaria para poder acomodar el elevado número de fosfatos (hasta seis) del inositido (**Figura 4.14**).





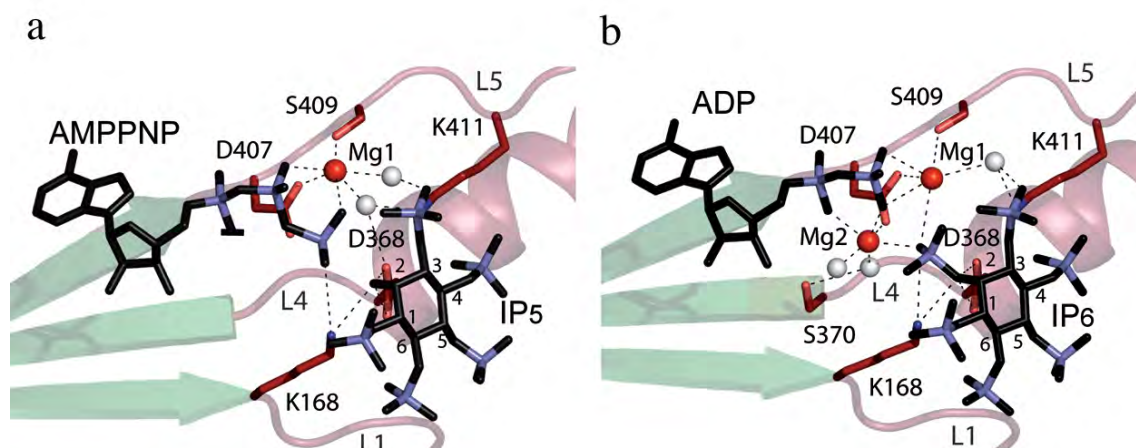
**Figura 4.14. Superficie electrostática de IP<sub>5</sub> 2-K.** Los colores representan la carga de las diferentes zonas de la superficie de la proteína, representando los tonos azules cargas positivas, los tonos blancos ausencia de carga y los tonos rojos carga negativa. El IP<sub>5</sub> se representa en color amarillo con los fosfatos resaltados en verde.

##### 4.1.5.3 La unión del metal en el centro activo

La enzima presenta iones  $Mg^{2+}$  en el centro activo; el complejo ternario con sustratos (IP<sub>5</sub>/AMPPNP) muestra un magnesio (Mg1), mientras que el complejo ternario con productos (IP<sub>6</sub> y ADP) muestra dos (Mg1 y Mg2). En cualquier caso, la presencia de magnesio parece estar ligada a la unión de nucleótido o al proceso de catálisis, ya que los complejos binarios que sólo presentan el inositido no lo incorporan.

En el primer caso, el ión Mg1 presenta coordinación octaédrica, cuyos ligandos son los fosfatos P $\gamma$  y P $\beta$  del ATP, Asp407, Ser409 y dos moléculas de agua (**Figura 4.15a**).

En el complejo ternario con ambos productos se encuentra un ión  $Mg^{2+}$  adicional (Mg2). En este complejo, tanto Mg1 como Mg2 están unidos al nucleótido y al inositido. El Mg1 está coordinado por los mismos elementos anteriores, siendo sustituido el P $\gamma$  por el P2 del IP<sub>6</sub> y una de las moléculas de agua por una interacción adicional con el P $\beta$ . Los ligandos del Mg2 son el P2 del IP<sub>6</sub>, P $\beta$  y P $\alpha$  del ADP, Asp407 y dos moléculas de agua, una de las cuales está coordinada por Ser370 (**Figura 4.15b**). Todos los metales del centro activo satisfacen la geometría esperada para estos iones. Su función parece ser la orientación del P $\gamma$  para el posterior ataque nucleófilo del inositido durante el proceso de catálisis.

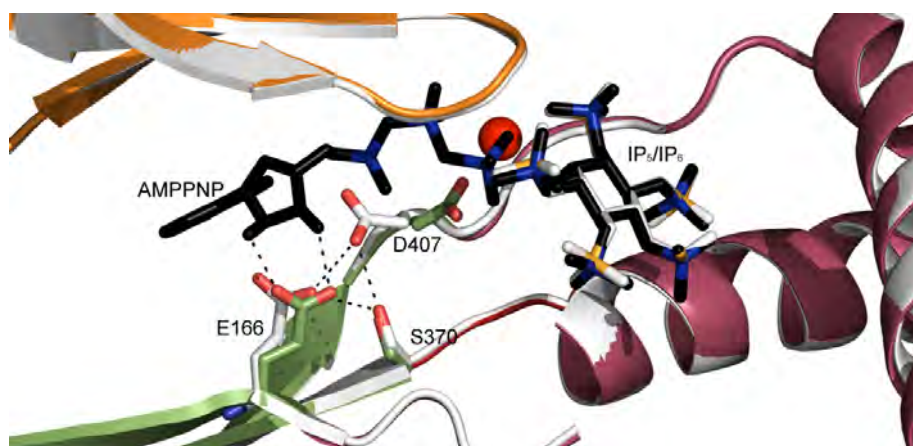


**Figura 4.15. Interacciones de los átomos de  $Mg^{2+}$  en el centro activo.** La estructura de la proteína se muestra mediante un esquema de cintas según el código de colores de la **Figura 4.7**. Los iones  $Mg^{2+}$  se representan mediante esferas en color rojo y las moléculas de agua mediante esferas de color blanco. Los ligandos se muestran en color negro, con los fosfatos en azul. a) Interacciones del Mg1 en el complejo ternario  $IP_5$  2-K/ $IP_5$ /AMPPNP. b) Interacciones del Mg1 y el Mg2 en el complejo ternario  $IP_5$  2-K/ $IP_5$ /ADP.

Es destacable que los iones  $Pb^{2+}$  determinados en el complejo  $IP_5$  2-K/ $IP_5$ /ADP/ $Pb^{2+}$  que permitió la resolución estructural, se localizan en posiciones equivalentes a las descritas para Mg1 y Mg2, mimetizando la interacción entre estos iones y la enzima.

#### 4.1.5.4 Cambios en el centro activo unidos a la entrada de nucleótido

El análisis de los complejos revela algunas diferencias locales en ciertos residuos del centro activo, consecuencia de la unión del nucleótido. El nucleótido provoca un cambio de conformación en la cadena lateral del Asp407 (L5) la cual coordina en este caso con el Mg1, dejando de interaccionar a su vez con Ser370 (**Figura 4.16**). Además, el nucleótido provoca cambios en la cadena lateral de Glu166. Este residuo se encuentra coordinado con Ser370 cuando el nucleótido está ausente, sin embargo, al unirse el nucleótido se rompe la coordinación anterior y Glu166 establece interacción con el nucleótido.



**Figura 4.16. Cambio de conformación del D407 debido a la presencia de AMPPNP.** Las estructuras de IP<sub>5</sub> 2-K/IP<sub>5</sub> e IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP fueron superpuestas para la realización de la imagen. IP<sub>5</sub> 2-K/IP<sub>5</sub> se representa mediante color blanco, mientras que IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP está representado según los colores empleados en la **Figura 4.7**. Los residuos correspondientes a cada complejo se representan empleando los respectivos colores.

### 4.1.5.5 El mecanismo catalítico de la IP<sub>5</sub> 2-K

La cristalización de los diferentes complejos de IP<sub>5</sub> 2-K ha permitido identificar los residuos que coordinan el P $\gamma$  y el 2-OH (antes de la reacción), así como del P $\beta$  y el P2 (después de la reacción), hecho que permite proponer un posible mecanismo de catálisis para la enzima. Básicamente, los residuos implicados provienen del bucle G, los bucles L1 (Lys168), L4 (Asp368) y L5 (Asp407 y Ser409, ambos a través del Mg1). El Asp368 parece crítico en el posicionamiento del grupo receptor del fosfato del anillo de inositol (2-OH), seleccionando el isómero que lo presenta en posición axial. Por otro lado, el P $\gamma$  estaría orientado adecuadamente vía interacciones con los residuos del bucle G, y el par Asp407/Ser409 (estas últimas interacciones a través del ión Mg1). Teniendo en cuenta la escasa distancia entre el P $\gamma$  donador y el 2-OH aceptor (distancia P $\gamma$  - 2-OH = 3.2 Å; P $\beta$ -P2 = 3.1 Å) es probable que la transferencia del fosfato se produzca “en línea” tal y como se ha propuesto para otras kinasas (Zhou *et al.*, 1998; Blaszczyk *et al.*, 2000). Finalmente, la Lys168 parece jugar un papel relevante en la neutralización de la carga negativa desarrollada en el estado de transición. El papel del segundo ión magnesio no está claro, si bien podría estar asociado a la estabilización de los productos del centro activo.

### 4.1.6 El sitio de unión a Zn<sup>2+</sup>

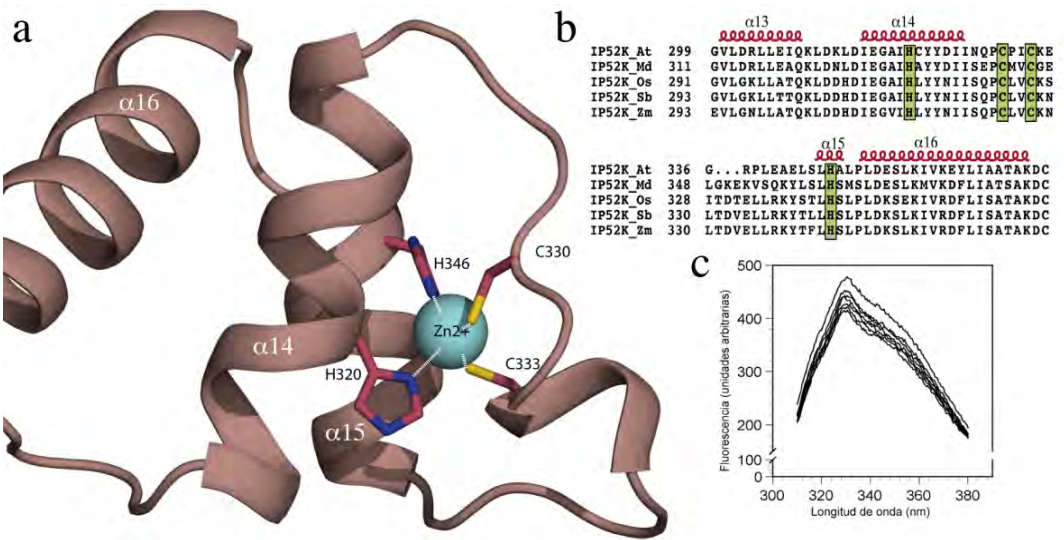
Como se mencionó anteriormente, las estructuras determinadas han revelado la presencia de un ión Zn<sup>2+</sup> en la estructura de IP<sub>5</sub> 2-K (Figuras 4.7 y 4.9). Las pruebas de espectrometría de emisión atómica por plasma acoplado inductivamente (ICP-OES)

confirmaron la presencia de 1 mol de  $Zn^{2+}$  por cada mol de proteína. El sitio de unión a  $Zn^{2+}$  se encuentra localizado en la región II de la región  $C_{IP}$  ( $C_{IP-II}$ ), concretamente en las hélices  $\alpha 14$  y  $\alpha 15$ . El  $Zn^{2+}$  está coordinado por dos cisteínas (Cys330, Cys333) y dos histidinas (His320 e His346) (**Figura 4.17a**). Esta zona está conservada únicamente en las  $IP_5$  2-K de plantas (**Figura 4.17b**).

**Tabla 4.5. Tasa de conversión de  $IP_5$ .** Estimación de la tasa de conversión de  $IP_5$  en  $IP_6$ , (nmol/min/ $\mu$ g de enzima) a 25 °C para la muestra control y la muestra tratada con 50 mM de EDTA. Cada ensayo consistió en 5 medidas independientes de actividad, y el ensayo se repitió en tres ocasiones.

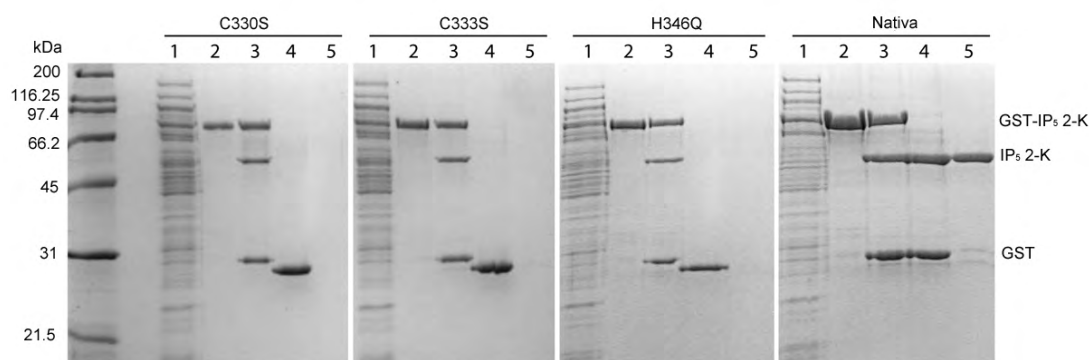
	Media
Control	8.9±0.7
Tratada con EDTA	10.3±1.8

Para estudiar la función del  $Zn^{2+}$  se realizaron diferentes experimentos. Por un lado, se midió la fluorescencia intrínseca de la proteína tras sucesivas adiciones de EDTA. Los sucesivos incrementos de la concentración de EDTA conllevaron una reducción de la fluorescencia de un 20% con respecto a la proteína nativa (**Figura 4.17c**), siendo esta reducción un proceso irreversible. Sin embargo no se observaron efectos en la actividad enzimática de la proteína (**Tabla 4.5**).



**Figura 4.17. Sitio de unión a  $Zn^{2+}$  de  $IP_5$  2-K.** a) Coordinación del  $Zn^{2+}$ . El  $Zn^{2+}$  se representa mediante una esfera de color azul claro y los residuos que lo coordinan en color cereza, con los átomos de nitrógeno y azufre en azul y amarillo respectivamente. b) Alineamiento entre diferentes secuencias de  $IP_5$  2-K de plantas mostrando la subregión  $C_{IP-II}$ : At: *A. thaliana*, Md: *Malus domestica*, Os: *Oryza sativa*, Sb: *Sorghum bicolor*, Zm: *Zea mays*. Los aminoácidos que forman el sitio de unión a  $Zn^{2+}$  están sombreados en verde. c) Espectro de emisión de fluorescencia de  $IP_5$  2-K. Cada curva se realizó a una concentración diferente de EDTA (0-2.5 mM) en incrementos de 0.5 mM. La sucesivas adiciones causan una disminución de la fluorescencia.

Por otra parte, se realizaron mutaciones de varios de los ligandos de este ión: C330S, C333S y H346Q. Todas las muestras mutantes fueron inestables en las condiciones en las cuales se produjo la proteína nativa, puesto que al cortar la etiqueta de fusión con TEV proteasa se degradaron por completo (**Figura 4.18**).



**Figura 4.18. Purificación de mutantes GST-IP<sub>5</sub> 2-K en el sitio de Zn<sup>2+</sup>.** Calle 1: lisado. Calle 2: GST-IP<sub>5</sub> 2-K unida a resina de GST. Calle 3: muestra tras 2 horas de incubación con proteasa TEV (sobrenadante + resina de GST). Calle 4: misma muestra que calle 3 tras 12 horas de incubación con proteasa TEV (sobrenadante + resina). Calle 5: sobrenadante tras 12 horas de incubación con proteasa TEV. Solo la proteína nativa se mantiene en el sobrenadante tras el corte.

El conjunto de estos resultados parece indicar que el Zn<sup>2+</sup> podría jugar un papel estructural, ayudando al correcto plegamiento de la enzima, aunque no participa directamente en la actividad catalítica.

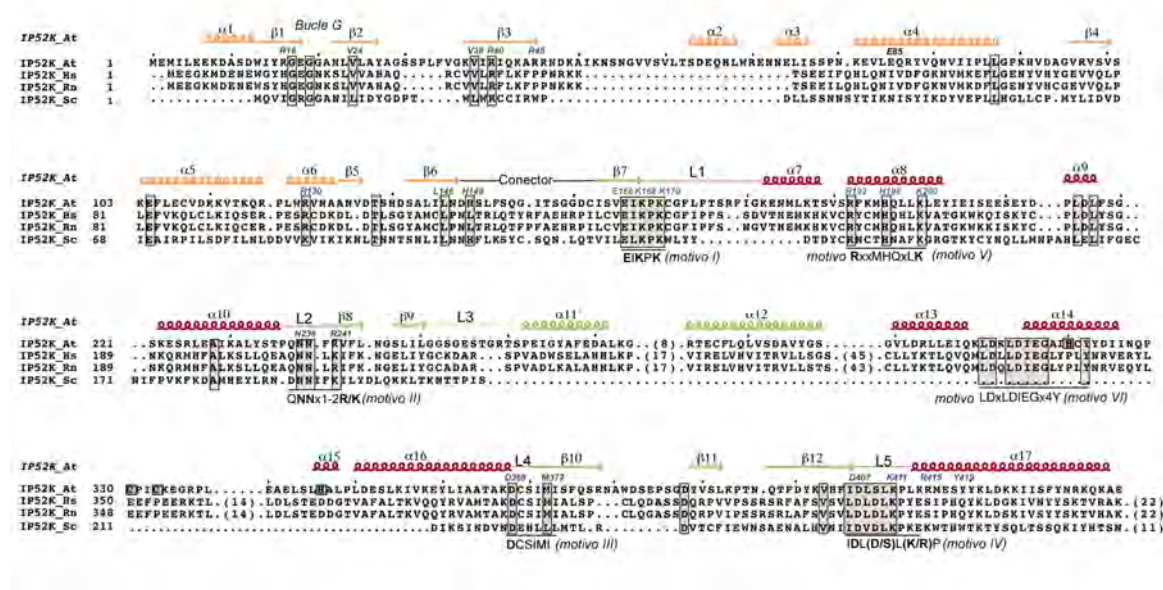
### 4.1.7 La IP<sub>5</sub> 2-K de *A. thaliana* como modelo para la familia de las IP<sub>5</sub> 2-K

Con el fin de profundizar en las similitudes y diferencias entre las IP<sub>5</sub> 2-K de diferentes especies se analizó un alineamiento de las diferentes secuencias (67 en total) de IP<sub>5</sub> 2-K encontradas en la base de datos PFAM (Punta *et al.*, 2012). Algunas de las secuencias más representativas se muestran en la **Figura 4.19**. La predicción de las estructuras secundarias así como los residuos conservados entre ellas, permiten proponer que la mayoría de elementos de estructura secundaria observados en la IP<sub>5</sub> 2-K de *A. thaliana* están conservados en las enzimas de diferentes especies. En concreto, las láminas  $\beta$  de ambos lóbulos, así como las mayoría de las hélices  $\alpha$  se encuentran conservados.

Sin embargo también pueden inferirse una serie de diferencias estructurales relevantes. En primer lugar, las hélices  $\alpha 2$  y  $\alpha 3$  del lóbulo N<sub>l</sub> parecen restringidas a enzimas de plantas. Por otro lado, las enzimas de mamíferos presentan dos inserciones adicionales

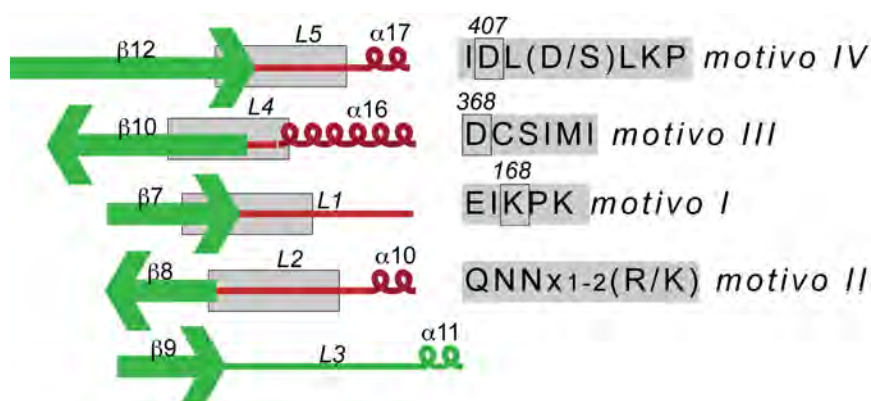


en la región C<sub>IP-II</sub>. La primera, de 45 aminoácidos de longitud, se encuentra localizada entre  $\alpha 12$  y  $\alpha 13$ , no mostrando una predicción de estructura secundaria clara. El segundo de estos insertos, de 14 aminoácidos de longitud, está localizado entre  $\alpha 14$  y  $\alpha 15$  y se encuentra en una zona equivalente al sitio de unión a Zn<sup>2+</sup>, cuya predicción de estructura secundaria sugiere la formación de una hélice  $\alpha$  adicional. Las enzimas de levaduras y otros hongos carecen de región C<sub>IP-II</sub>.



**Figura 4.19. Alineamiento de IP<sub>5</sub> 2-Ks de diferentes especies.** At: *A. thaliana*, Hs: *H. sapiens*, Rn: *Rattus norvegicus*, Sc: *S. cerevisiae*. Los elementos de estructura secundaria se muestran en la parte superior, coloreados según el código de colores mostrado en la **Figura 4.7**. Los residuos recuadrados en color verde corresponden al sitio de unión a Zn<sup>2+</sup>, mientras que los recuadros en color marrón indican residuos conservados entre las diferentes especies. Los residuos implicados en la unión de nucleótido e inositol se representan en la zona superior con letra en color verde y azul respectivamente. Los diferentes motivos se encuentran señalados en la parte inferior.

Por otro lado, el alineamiento entre las diferentes secuencias de IP<sub>5</sub> 2-K revela que existen seis motivos conservados en las IP<sub>5</sub> 2-K de diferentes especies. Según la estructura, cuatro de estos motivos (motivos I-IV) están situados en los bucles L1 (motivo I; <sup>166</sup>EIKPK), L2 (motivo II; <sup>237</sup>QNNx<sub>1-2</sub>(R/K)), L4 (motivo III; <sup>368</sup>DCSIMI) y L5 (motivo IV; <sup>406</sup>IDL(D/S)LKP). La mayoría de los residuos importantes para la actividad y unión del inositol se encuentran en estos motivos (**Figura 4.20**), como por ejemplo Lys168, Asp368 o Asp407. En cuanto al bucle L3, no conservado, es probable que esté implicado en aspectos particulares para cada especie, tales como la cinética de reacción y/o la modulación de la actividad catalítica.



**Figura 4.20. Disposición de los bucles que albergan los diferentes motivos conservados implicados en la actividad enzimática.** Los diferentes elementos de estructura secundaria son mostrados empleando el mismo código de colores que en la **Figura 4.7** y siguientes. Los segmentos grises indican la posición exacta de los motivos conservados (cuya secuencia se muestra a la derecha). Los residuos relevantes para la catálisis se muestran recuadrados.

Además de estos cuatro motivos conservados, existen dos motivos adicionales: el motivo <sup>192</sup>RxxMHQxLK (motivo V), implicado en el reconocimiento del anillo de inositol y el motivo <sup>310</sup>LDxLDIEGxxxxY (motivo VI), que parece jugar un papel importante en el mantenimiento de la estructura de la región C<sub>IP-II</sub>. Todas las especies conservan estos motivos excepto los hongos, que carecen del motivo VI, en concordancia con la ausencia de la región C<sub>IP-II</sub> en los mismos.

#### 4.1.8 La familia de la IP<sub>5</sub> 2-K y su relación con la familia de las IPKs

##### 4.1.8.1 La IP<sub>5</sub> 2-K conserva el plegamiento de la familia de IPKs

La resolución estructural de la IP<sub>5</sub> 2-K permitió identificar que esta familia conserva ciertos elementos del plegamiento de la familia de las IPKs. Una búsqueda de homólogos estructurales empleando el servidor DALI (Holm y Rosenstrom, 2010) reveló que el homólogo estructural más cercano es una IPK, en concreto la IP<sub>3</sub>-3K C (rmsd = 5.4 Å para 154 residuos, incluyendo el lóbulo N<sub>t</sub> y la lámina β del C<sub>t</sub>).

En el N<sub>t</sub> se conserva el plegamiento general del lóbulo (principalmente lámina β y hélice α4), así como el patrón de reconocimiento del nucleótido. En cuanto al lóbulo C-terminal, éste también presenta cierta homología estructural con la familia de enzimas IPKs. En concreto, esta homología se centra en la lámina β y hélices α11 y α12,

mientras que la región de unión a IP es completamente diferente. De hecho, no se encontró ninguna proteína que presentara homología estructural con la región C<sub>IP</sub>, por tanto, éste muestra un plegamiento específico de la familia de IP<sub>5</sub> 2-Ks.

Se puede concluir que la IP<sub>5</sub> 2-K pertenece a la familia de las inositol polifosfato kinasas (IPKs), al contrario de lo establecido hasta la fecha, que clasificaba estas enzimas como una familia diferenciada y separada de las IPKs (Cheek *et al.*, 2005). Sin embargo es el miembro más lejano de esta familia de las IPKs, mostrando una región de unión a inositol de gran tamaño y completamente diferente.

Un alineamiento de secuencia realizado entre IP<sub>5</sub> 2-K y diferentes miembros de la familia de las IPKs (**Figura 4.21a**) representa cómo la mayor parte del lóbulo N<sub>i</sub>, así como parte del lóbulo C<sub>i</sub> (lamina β, y hélices α11 y α12) se encuentran conservadas entre IP<sub>5</sub> 2-K e IPKs.

#### 4.1.8.2 Inserciones en las IP<sub>5</sub> 2-K. Comparación con inserciones en las IPKs

Uno de los elementos distintivos entre las IPKs es la presencia de inserciones a lo largo de su secuencia. En base a los alineamientos realizados entre las distintas IPKs, se han localizado dos regiones propensas a presentar inserciones, señaladas como inserto I e inserto II (**Figura 4.21a y b**). Estos insertos se han relacionado con funciones de regulación. Por ejemplo en algunas IPKs (como en IP<sub>6</sub>-K 2) el inserto I ha sido relacionado con la interacción de esta proteína con Hsp90 (Chakraborty *et al.*, 2008) y p53 (Koldobskiy *et al.*, 2010), con lo que la función de este inserto podría ser variable según la enzima. En el caso de la IP<sub>5</sub> 2-K, no se ha reportado ninguna interacción de esta zona con otras proteínas. Por su localización, el inserto I podría estar relacionado con el reconocimiento de sustrato, tal y como se ha observado en IP<sub>5</sub> 2-K (contiene α6, y R160).

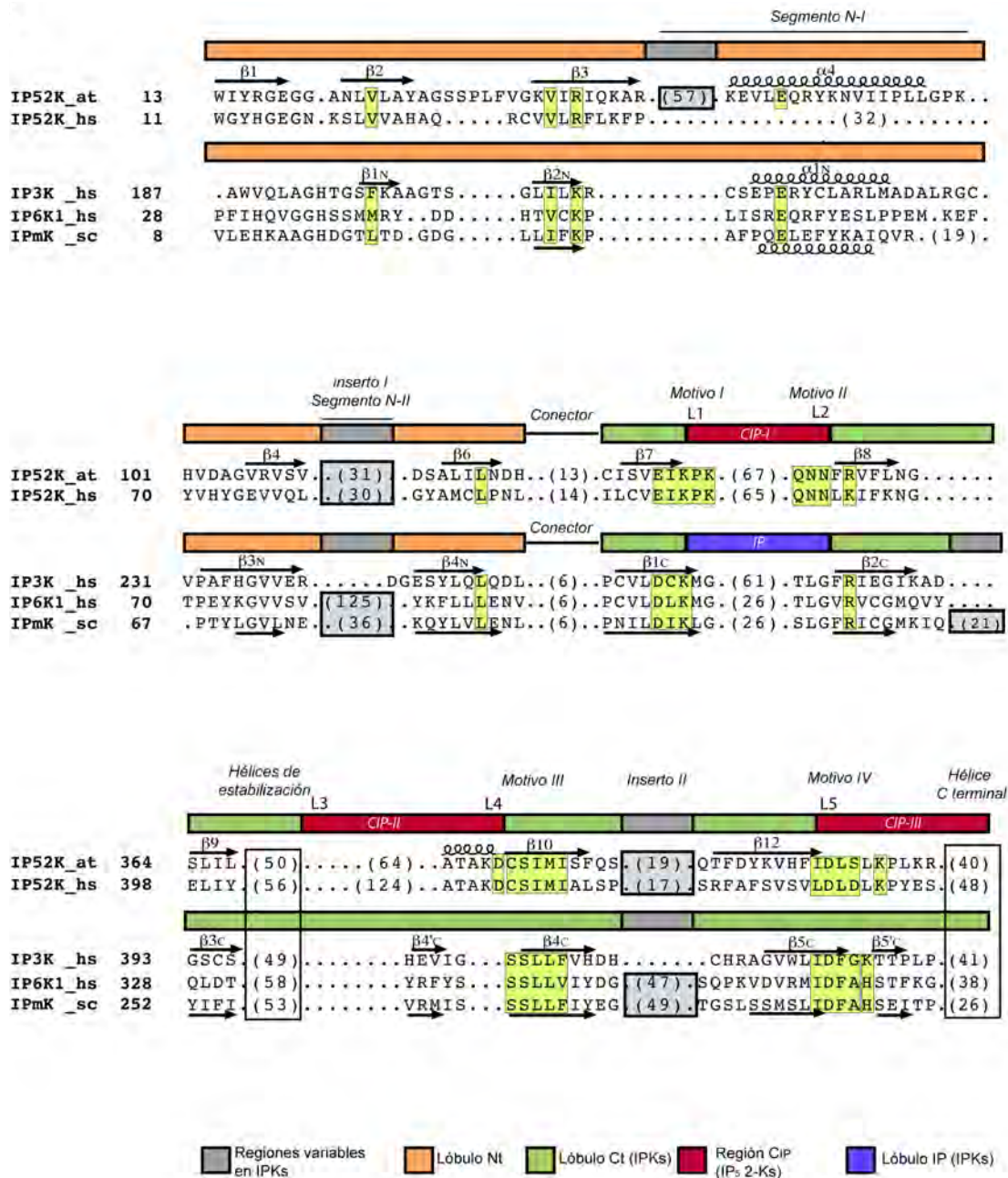
En cuanto al inserto II, éste ha sido también relacionado con interacciones proteína-proteína (como por ejemplo en el caso de IPMK) (Messenguy *et al.*, 1991), aunque este hecho aún no ha sido corroborado (Caddick *et al.*, 2007). En el caso de IP<sub>5</sub> 2-K, este inserto no se ha relacionado aún con ninguna interacción proteína-proteína.

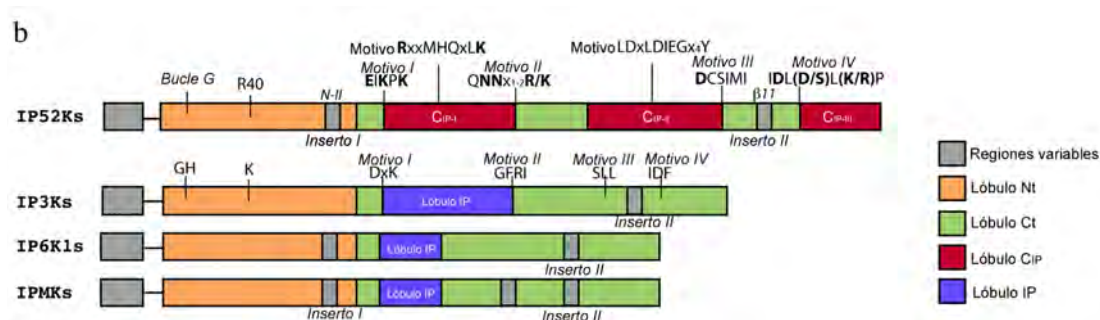
Las IP<sub>3</sub> 3-Ks presentan además dominios reguladores, situados en el extremo N-terminal de su secuencia (Xia y Yang, 2005). En el caso de las IP<sub>5</sub> 2-Ks sólo la enzima de *Schizosaccharomyces pombe* presenta un fragmento que podría cumplir dicha función, aunque no ha sido determinada ninguna función concreta para este fragmento hasta la fecha (Ives *et al.*, 2000).



#### 4. Resultados y discusión

En resumen los insertos y regiones reguladoras son elementos muy variables dentro de la familia de las IPKs tanto en tamaño como en función. Gracias a la estructura de la IP<sub>5</sub> 2-K se ha podido determinar por primera vez que el inserto I podría estar relacionado con la unión del sustrato en esta familia de proteínas.





**Figura 4.21. Organización de la familia IPK.** a) Alineamiento entre los miembros de la familia IPKs. IP52K\_at: IP<sub>5</sub> 2-K de *A. thaliana*, IP52K\_hs: IP<sub>5</sub> 2-K de *H. sapiens*, IP3K\_hs: IP<sub>3</sub> 3-K A de *H. sapiens*, IP6K1\_hs: IP<sub>6</sub>-K 1 de *H. sapiens*, IPmK\_sc: IPmK de *S. cerevisiae*. Las diferentes regiones se encuentran señaladas mediante los rectángulos de la parte superior del dibujo, según los colores explicados en la propia figura. Los elementos de estructura secundaria se encuentran señalados encima del propio alineamiento, según las nomenclaturas empleadas para IP<sub>5</sub> 2-K (en esta tesis) e IP<sub>3</sub> 3-K (González *et al.*, 2004). Los residuos conservados se encuentran recuadrados en color verde. b) Esquema de la organización de dominios de cada uno de los miembros de la familia IPK. Los diferentes elementos clave de cada una de las enzimas se encuentran señalados en la parte superior del esquema de cada enzima.

#### 4.1.8.3 El lóbulo de unión a inositido: un elemento de divergencia entre los diferentes miembros de la familia IPK

Uno de los puntos de mayor divergencia entre las diferentes proteínas pertenecientes a la familia IPK es la región de unión a inositido (región C<sub>IP</sub> en la IP<sub>5</sub> 2-K, lóbulo IP para el resto de IPKs). En todos los casos se trata de una región formada exclusivamente por hélices  $\alpha$ , pero su tamaño varía entre los diferentes miembros de la familia. En el caso de la región C<sub>IP</sub> de IP<sub>5</sub> 2-K su extensión abarca 190 residuos (un total de 9 hélices  $\alpha$ ), siendo de un tamaño más modesto en el caso de la IP<sub>3</sub> 3-K (60 residuos, 4 hélices  $\alpha$ ), y muy reducido en el caso de las IPMKs (30 residuos, 2 hélices  $\alpha$ ) (**Figura 4.22**). El plegamiento de los distintos lóbulos es totalmente diferente.

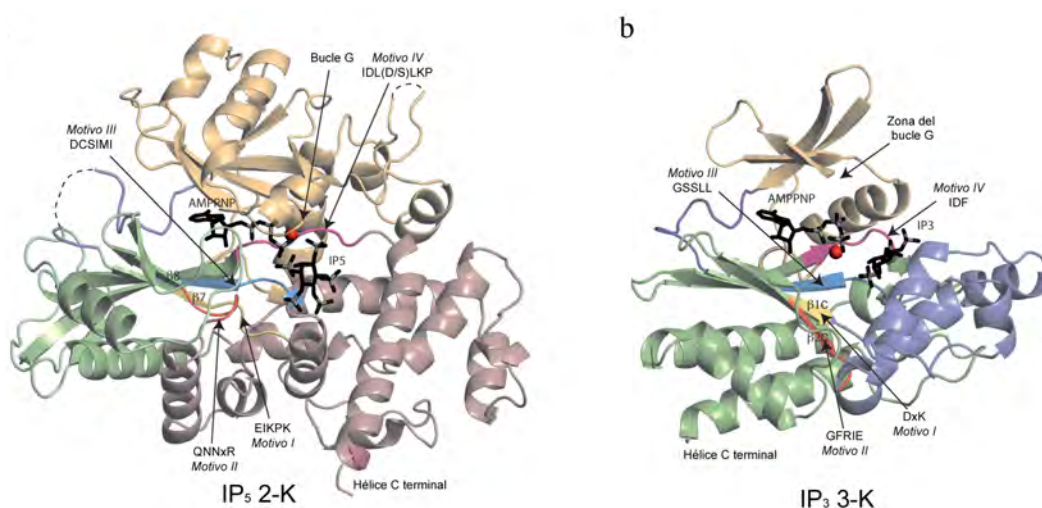
Considerando que IP<sub>5</sub> 2-K es la única familia de enzimas conocida que cataliza la fosforilación de una posición axial del inositol, es interesante comparar el modo de reconocimiento del inositido con el de otras enzimas la familia de IPKs. Con este fin se ha realizado una comparación entre el centro activo de IP<sub>5</sub> 2-K e IP<sub>3</sub> 3-K. La superposición del sitio de unión al nucleótido de ambas estructuras revela que las posiciones fosforilables (posición 2 axial para IP<sub>5</sub> 2-K y 3 ecuatorial para IP<sub>3</sub> 3-K) presentan la misma orientación en relación al nucleótido, lo que produce que la disposición de ambos anillos sea perpendicular entre sí (**Figura 4.23a y b**). Esto conlleva que los bolsillos de unión de inositol sean muy diferentes en ambos casos y que las interacciones no estén conservadas en ambas enzimas. Todos los demás

elementos de reconocimiento son diferentes, entre los cuales cabe destacar un residuo específico y conservado solo en las IP<sub>5</sub> 2-K, Asp368, encargado de la selección del isómero *myo*-inositol (es decir, que el 2-OH se encuentre en posición axial).



**Figura 4.22. Lóbulos de unión a IP de cada uno de los miembros de la familia IPK.** El lóbulo de unión a inositido de cada molécula se representa mediante color magenta, estando el resto de la molécula representada en color blanco. Los ligandos de cada estructura se representan en color negro. a) Lóbulo de unión a inositido correspondiente a IPMK de *S. cerevisiae* (Holmes y Jogl, 2006). b) Lóbulo de unión a inositido de IP<sub>3</sub>-3K A de *H. sapiens* (González *et al.*, 2004; Miller y Hurley, 2004). c) Lóbulo de unión a inositido de IP<sub>5</sub> 2-K de *A. thaliana*.

En conclusión, no hay homología ni en el lóbulo de unión ni en el modo de reconocimiento del inositido en IP<sub>5</sub> 2-K con el resto de la familia de IPKs. Puede decirse que las únicas características comunes entre los centros activos de ambas enzimas son el fuerte carácter positivo de éste, el reconocimiento del nucleótido, y se verá a continuación, los residuos que intervienen en la catálisis.



**Figura 4.23. IP<sub>5</sub> 2-K vs IP<sub>3</sub> 3-K A.** a) Estructura general de IP<sub>5</sub> 2-K. El código de colores para los lóbulos y regiones de la enzima es el mismo que el empleado en la **Figura 4.7**. Se señala la ubicación de los diferentes motivos de la enzima mediante el uso de colores: motivo I: amarillo, motivo II: naranja, motivo III: azul, motivo IV: rosa. Los ligandos están representados en color negro. b) Estructura general de IP<sub>3</sub> 3-K A. El código de colores es el mismo que el explicado para la figura a). En este caso la región de unión al inositido se representa en color azul oscuro.

#### 4.1.8.4 Los residuos catalíticos: conservación a través de la familia IPK

Pese a las diferencias expuestas en el modo de reconocimiento del inositido pueden encontrarse residuos equivalentes a los propuestos como catalíticos en IP<sub>5</sub> 2-K también en otras IPKs. Estos residuos son la Lys168, que parece neutralizar el estado de transición, y el Asp407, que orienta el P $\gamma$ . Estos residuos se encuentran conservados en las otras tres familias de IPKs (**Figura 4.21**). Asimismo, el bucle G se encuentra presente en todas estas enzimas.

Sin embargo, la IP<sub>5</sub> 2-K presenta residuos específicos, no conservados con otras IPKs, como el Asp368 o la Ser409, que podrían estar implicados en catálisis, ya que coordinan el 2-OH receptor y el magnesio respectivamente.

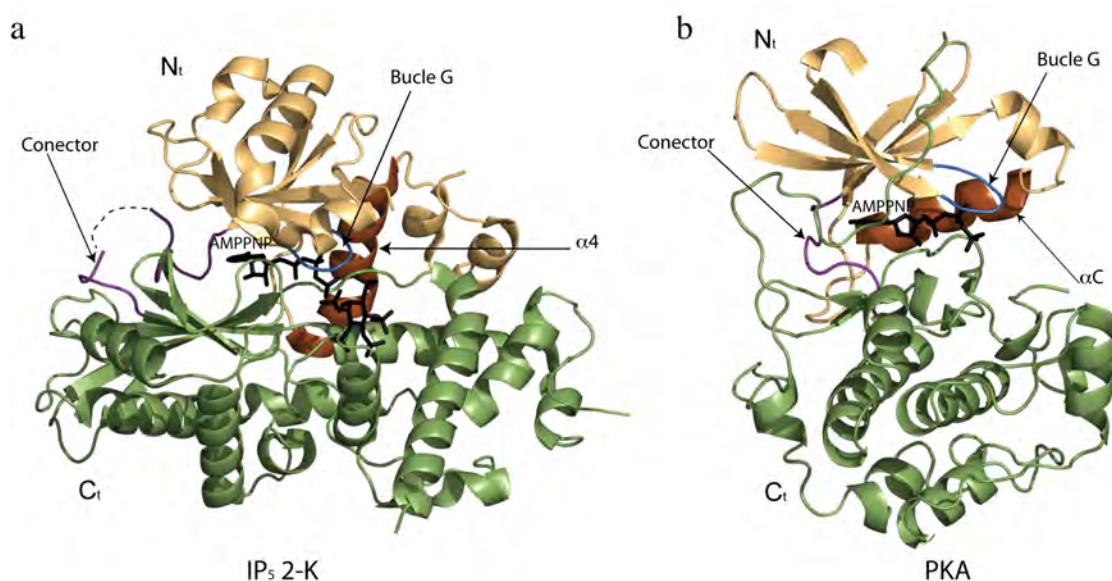
Como se explicó en anteriores secciones, los residuos clave para la actividad de IP<sub>5</sub> 2-K están agrupados en diferentes motivos conservados (motivos I-IV). De ellos, los motivos I y IV engloban los residuos catalíticos citados arriba, estando conservados en todas las IPKs (**Figura 4.21**). Los motivos II y III agrupan residuos que son importantes para la coordinación del inositol. La secuencia de estos motivos varía de unas IPKs a otras, aunque no obstante se encuentran en la misma zona estructural (**Figura 4.23a y b**). Los motivos V y VI son exclusivos de IP<sub>5</sub> 2-K, no existiendo en otra familia de enzimas.

#### 4.1.9 La IP<sub>5</sub> 2-K y su relación con las proteínas kinasas (PKs)

La resolución estructural de la IP<sub>5</sub> 2-K permitió constatar que, al igual que sucede con la familia de las IPKs, ésta conserva ciertos elementos del plegamiento de las proteína kinasas (PKs). Uno de los homólogos estructurales más cercanos de la IP<sub>5</sub> 2-K encontrados mediante DALI es la proteína kinasa HASPIN (rmsd = 4.2 Å para 108 residuos).

La homología estructural con las proteínas kinasas se centra en el N<sub>t</sub>. En particular, se conserva la lámina  $\beta$  y la  $\alpha 4$  (equivalente a la llamada  $\alpha C$  de proteína kinasas). Otro de los elementos estructurales conservados es el conector entre los lóbulos N y C-terminal (**Figura 4.24**).



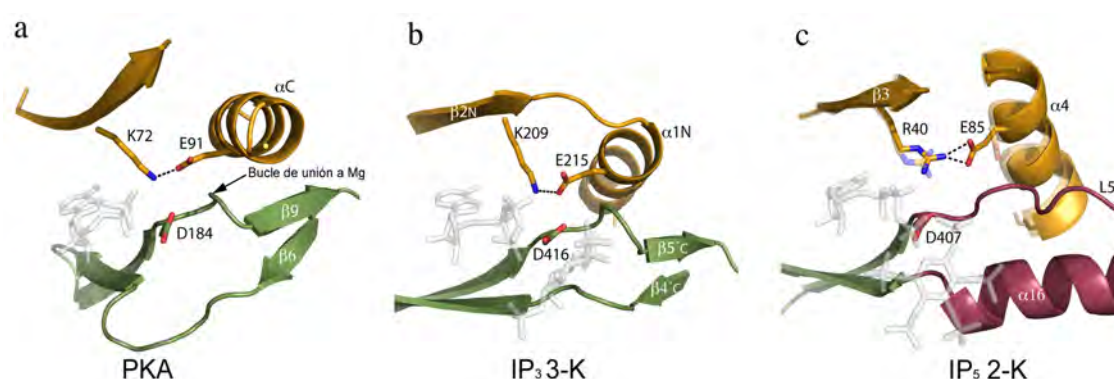


**Figura 4.24. IP<sub>5</sub> 2-K y proteína kinasa A.** a) IP<sub>5</sub> 2-K. El dominio N<sub>t</sub> está representado en color amarillo, el dominio C<sub>t</sub> en verde y el conector entre ambos dominios en color violeta. El bucle G está destacado en color azul y la hélice α4 en color marrón. b) Estructura de PKA. El código de colores es el mismo que se ha empleado en la figura a). En marrón se representa la αC típica de las PKs.

El sitio de unión al nucleótido se encuentra conservado en ambas enzimas, incluyendo el bucle G, un elemento crucial caracterizado originalmente en PKs. Además, los residuos catalíticos descritos en la IP<sub>5</sub> 2-K, los cuales se conservan en la familia de las IPKs, se encuentran conservados también en la familia de las PKs. Estos son la Lys168 (Lys168 en PKA) y el Asp407 (Asp184 en PKA).

Las proteínas kinasas presentan una serie de elementos cuya conformación marca su estado activo o inactivo. Uno de los elementos típicos de una PK activa es la formación de un puente salino (Lys72/Glu91 en el caso de la PKA), asociado a la correcta orientación de una hélice del N<sub>t</sub> denominada αC (**Figura 4.24**). Esta característica también se encuentra presente en IPKs, sin embargo no está presente en IP<sub>5</sub> 2-K (**Figura 4.25**). Aunque sí puede observarse un puente salino equivalente (puente entre Arg40 y Glu85), éste solo se observa en una de las dos moléculas de la unidad asimétrica, no afectando su eliminación a la actividad de la proteína (ver siguiente sección).

Otro elemento distintivo de la conformación activa de las PKs e IPKs es la formación de una pequeña lámina β (**Figura 4.25a y b**) encargada de posicionar de forma adecuada el bucle de unión a Mg<sup>2+</sup>. Sin embargo, este también se encuentra ausente en la IP<sub>5</sub> 2-K (**Figura 4.25c**). La ausencia de estos elementos que marcan la conformación activa de las PKs y las IPKs indica que los mecanismos de activación de IP<sub>5</sub> 2-K deben ser diferentes a los de estos dos grupos de enzimas.



**Figura 4.25. Los elementos típicos de la conformación activa de PKs están ausentes en la IP<sub>5</sub> 2-K.** El código de colores se corresponde con el empleado en la **Figura 4.7**, estando los diferentes sustratos representados en color blanco. a) Puente salino en PKA (Lys72-Asp91), asociado a una orientación determinada de  $\alpha C$  que determina una conformación activa. Una pequeña lámina  $\beta$  ( $\beta 6$ - $\beta 9$ ) se forma para orientar el bucle de unión a magnesio. b) Los elementos mostrados en a) se encuentran conservados en IP<sub>3</sub>-3K A, formándose la lámina  $\beta$  entre  $\beta 4'C$  y  $\beta 5'C$ . c) Estos elementos no se conservan en el caso de IP<sub>5</sub> 2-K. Pese a que en una de las moléculas de la unidad asimétrica puede observarse un puente salino equivalente al que muestra la conformación activa de PKA la orientación de la hélice  $\alpha$  equivalente a  $\alpha C$  no es la misma. La presencia de  $\alpha 16$  parece evitar que exista una lámina  $\beta$  equivalente a la observada en PKA.

#### 4.1.10 Análisis mediante mutagénesis dirigida de la IP<sub>5</sub> 2-K

**Tabla 4.6. Actividad de los diferentes mutantes de IP<sub>5</sub> 2-K.** ND: actividad no detectable. Los valores entre paréntesis indican medidas de actividad realizadas a una concentración de proteína 10 veces superior a la empleada para el primer ensayo. Se realizaron medidas de  $K_m$  (para ATP) y  $V_{max}$  de la enzima nativa.  $K_m = 38 \mu M$  y  $V_{max} = 1080 \text{ pmol/min}/\mu g$ . Se midieron también estos valores para los mutantes R40V, E85A y N238A, obteniéndose valores de  $K_m$  2-3 veces superiores a los obtenidos para la enzima nativa, y descensos de la  $V_{max}$  de menos de un 20%.

	Actividad
Nativa	100
R40V	56.06
E85A	124.24
R130I	6.06 (9.09)
K168A	ND (2.27)
K168N	ND (4.54)
K170S	5.30 (3.78)
N238A	85.60
D368A	ND (ND)
D407A	3.03 (7.57)
K411A	ND (ND)

Con el objetivo de estudiar el papel de los diferentes residuos observados en la actividad de la enzima (Figuras 4.10, 4.11 y 4.12) se obtuvieron diferentes mutantes, cuya actividad fue caracterizada. La actividad de cada uno de ellos con respecto a la proteína nativa puede verse en la **Tabla 4.6**.

El mutante R40V mostró una disminución a la mitad de su actividad, probablemente debido a su incapacidad a la hora de coordinar los fosfatos  $P\alpha$  y  $P\beta$  del nucleótido. Esto se ve reforzado por el aumento en la  $K_m$  para la unión de ATP observado en este mutante en una proporción de 2-3 veces con respecto a la enzima no mutante. Como era esperable, los mutantes de los residuos catalíticos no mostraron actividad (K168A, K168N) o mostraron una actividad residual (D407A), apoyando su papel esencial en el proceso de catálisis. La completa inactividad del mutante de Arg130 (R130I) señala la importancia de que el P1 esté coordinado, así como del papel crucial que juega el lóbulo  $N_l$  para la actividad enzimática. Además este residuo interacciona con el bucle L3, estabilizando su posición, con lo que podría estar ayudando a mantener la adecuada conformación de la enzima para unir el inositol. Otro de los residuos claves para la coordinación del inositol es la Lys170, cuya disminución de actividad señala la importancia de la coordinación de los fosfatos P5 y P6. No obstante, la mutación N238A (residuo localizado en el bucle L2 que también coordina P6) tuvo efecto modesto en cuanto a disminución de la actividad, aumentando sin embargo su  $K_m$  para el ATP en 2-3 veces con respecto a la enzima nativa. El escaso efecto en la actividad de esta mutación puede deberse a la presencia de otros aminoácidos (Lys170, Lys200) que también coordinan dicho grupo fosfato. Sin embargo, la disminución de la afinidad por el ATP puede deberse a una disminución en la capacidad de interacción entre el L2 (donde se localiza el motivo II) y el bucle L3, lo cual podría propiciar que el bolsillo de unión a ATP no se cierre adecuadamente. Por último, hay otro par de residuos clave, cuya mutación conlleva la inactivación total de la enzima: Asp368 (coordina el 2-OH del inositido) y Lys411 (que coordina dos de los fosfatos del anillo, P3 y P4). El mutante E85A (residuo equivalente a Glu91 en la PKA) no mostró ninguna disminución en la actividad, confirmando por tanto que el puente Glu85-Arg40, equivalente al puente Glu91-Lys72 de las PKs, no tiene ninguna influencia en la actividad de la enzima y no parece estar implicado en su activación.

### 4.1.11 Conclusión

El trabajo expuesto hasta ahora muestra, por primera vez, la caracterización estructural de una  $IP_5$  2-K. La  $IP_5$  2-K de *A. thaliana* presenta una estructura  $\alpha+\beta$  compuesta de dos lóbulos N y C-terminal, unidos por un bucle de conexión. El sitio de unión a ATP, muy similar al que puede encontrarse en las proteínas kinasas, se encuentra situado en la interfase formada por las láminas  $\beta$  de los dos lóbulos N y C-terminal, estando residuos de ambos lóbulos implicados en la unión del compuesto. El

sitio de unión a inositido está principalmente conformado por una región  $\alpha$  helicoidal situada en el lóbulo C-terminal, denominada lóbulo C<sub>IP</sub>. La mayoría de residuos fundamentales para la unión de inositido y la catálisis provienen de esta región, aunque el lóbulo N-terminal también aporta algunos residuos clave. La relevancia de muchos de estos residuos ha sido estudiada mediante mutagénesis dirigida y ensayos de actividad. La proteína presenta además un sitio de unión a  $\text{Zn}^{2+}$ , que es probablemente un elemento clave para el correcto plegamiento de la enzima. Gracias al trabajo expuesto puede decirse que la IP<sub>5</sub> 2-K pertenece a la familia de las IPKs, si bien es el miembro más lejano de esta familia.

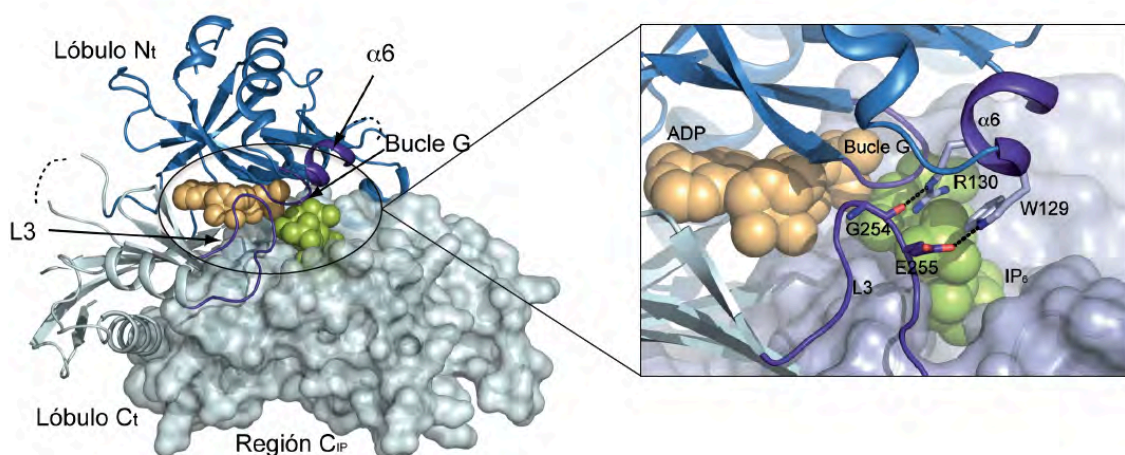
Pese al gran avance que supone el trabajo realizado, no se ha logrado caracterizar la estructura de IP<sub>5</sub> 2-K en su forma libre, desconociéndose los posibles cambios conformacionales de la enzima durante el ciclo catalítico. La Sección 4.2 tratará sobre los resultados obtenidos sobre éstas y otras cuestiones pendientes durante el desarrollo de este trabajo.



## 4.2 Estructura de IP<sub>5</sub> 2-K<sub>W129A</sub> de *A. thaliana*: apoenzima y complejos binarios con AMPPNP y ADP

### 4.2.1 Introducción

Pese a los esfuerzos realizados para cristalizar la apoenzima, en ningún momento se obtuvieron cristales de IP<sub>5</sub> 2-K en ausencia de inositido. Las diferentes estructuras obtenidas en presencia de IP se analizaron en profundidad con el fin de entender por qué éste es necesario para la cristalización de la enzima nativa.



**Figura 4.26. Estructura de IP<sub>5</sub> 2-K y cierre del centro activo.** A la derecha se muestra el complejo entre IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP. La enzima se encuentra representada mediante un modelo de cintas (lóbulo N<sub>t</sub> en color azul y C<sub>t</sub> en color azul claro), con la región C<sub>IP</sub> representada en forma de superficie. La hélice  $\alpha 6$ , el bucle L3 y el bucle G se encuentran destacados en violeta. El ADP se muestra mediante esferas de color amarillo anaranjado y el IP<sub>6</sub> mediante esferas de color verde. A la izquierda se observa la zona del centro activo ampliada, detallándose la interacción producida entre residuos de los elementos L3 y  $\alpha 6$ .

Como se ha mencionado anteriormente, el centro activo se encuentra cerrado mediante la interacción de la  $\alpha 6$  (Leu128-Ala133) con el bucle L3 (Gly251-Ser261), en concreto, mediante las interacciones Arg130-Gly254 y Trp129-Glu255 (**Figura 4.26**). La estructura cristalizada en presencia de inositido sugiere que estas interacciones podrían actuar como una especie de “cremallera” cuya apertura y cierre podría estar regulando la entrada y salida de los sustratos y productos. Es probable, por lo tanto, que ambos elementos,  $\alpha 6$  y L3, se encuentren implicados en un posible cambio conformacional.

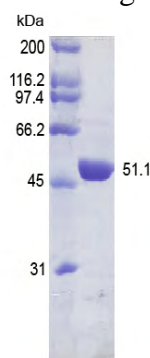
Con el fin de explorar esta posibilidad, se produjeron una serie de mutantes en los residuos de esta cremallera (ver Sección 4.2.8). Con estas mutaciones se pretendía

obtener muestras más susceptibles de ser cristalizadas en ausencia de inositido, así como analizar el papel de ciertos residuos en la enzima. Uno de estos mutantes, el mutante W129A ( $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ ) permitió conseguir este objetivo.

### 4.2.2 Obtención de muestras de $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ para estudios de cristalización

La expresión y purificación de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$  se realizó empleando el mismo protocolo que para la proteína nativa, partiendo también de 4 L de cultivo. Se obtuvieron resultados muy similares a los obtenidos con la proteína nativa en cuanto a grado de pureza y puntos de elución en cada paso del proceso de purificación. La elución de la proteína en la columna de heparina se realizó empleando una mezcla de 59% de tampón A y 41% de tampón B (390 mM de NaCl). En cuanto la cromatografía de exclusión molecular, la proteína eluía en el mismo punto que la proteína nativa (mL 77).

Pese a la mutación introducida, la cantidad de proteína obtenida fue similar a la de la enzima nativa, siendo de aproximadamente 12 mg/L de cultivo (**Figura 4.27**).

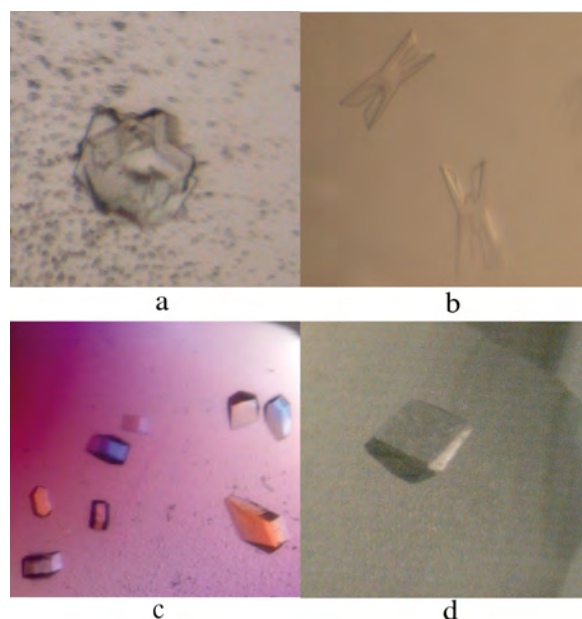


**Figura 4.27. SDS-PAGE de la muestra final de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ .** La banda se corresponde con la muestra del mutante W129A tras el proceso completo de purificación.

### 4.2.3 Cristalización de $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$

#### 4.2.3.1 Cristalización de la apoenzima

Del total de las 576 condiciones ensayadas, sólo se obtuvieron cristales en una condición (**Figura 4.28a**). Estas condiciones fueron ajustadas (**Figura 4.28b**) y la calidad de los cristales mejorada mediante sembrado (**Figura 4.28c**), técnica que también ayudó a que el tiempo de obtención de cristales disminuyese de dos semanas (condiciones iniciales) a 5 días. Las condiciones finales en las que se obtuvieron estos cristales fueron 29% (m/v) PEG 4000, 0.2 M  $\text{LiSO}_4$ , 0.1 M Tris-HCl pH 8.0.



**Figura 4.28. Cristalización de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$  e  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /AMPPNP.** a) Primeros cristales de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$  obtenidos (30% (m/v) PEG 4000, 0.2 M  $\text{LiSO}_4$ , 0.1 M Tris-HCl pH 8.5). b) Cristales de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$  obtenidos tras el ajuste de condiciones iniciales (32–34% (m/v) PEG 4000, 0.2 M  $\text{LiSO}_4$ , 0.1 M Tris-HCl pH 8.0). c) Cristales finales de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$  obtenidos mediante sembrado en 29% (m/v) PEG 4000, 0.2 M  $\text{LiSO}_4$ , 0.1 M Tris-HCl pH 8. d) Cristal de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /AMPPNP (obtenido mediante sembrado en 0.5 M  $(\text{NH}_4)_2\text{SO}_4$ , 0.1 M de citrato sódico pH 5.6 y 0.8 M  $\text{LiSO}_4$ ).

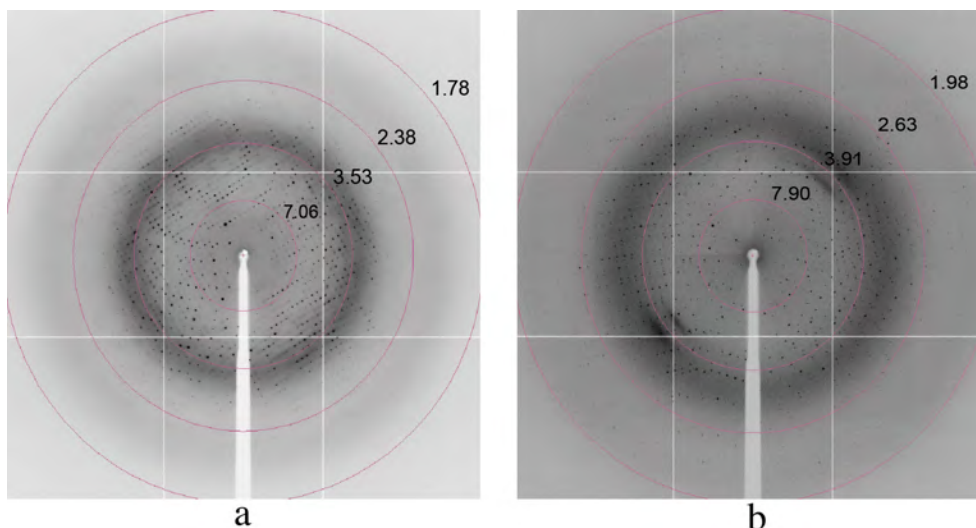
#### 4.2.3.2 Cristalización de los complejos $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ /AMPPNP e $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ /ADP

No se lograron obtener cristales de estos complejos en las mismas condiciones que la apoenzima, ni tampoco en las condiciones de la enzima nativa. Se buscaron nuevas condiciones de cristalización, que tras optimizar y usar la técnica de sembrado dieron lugar a los cristales que pueden apreciarse en la **Figura 4.28d**.

#### 4.2.4 Procesado de datos, resolución estructural y construcción de los modelos $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ , $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ /AMPPNP e $\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ /ADP

Se obtuvieron conjuntos de datos de difracción de rayos X de buena calidad (**Figura 4.29a y b**) en todos los casos. El procesamiento de los datos de difracción mostró que tanto los cristales de la apoenzima como los de los complejos binarios con el nucleótido pertenecen al grupo espacial  $\text{P2}_12_12_1$ . Comparando la celdilla de ambos puede apreciarse una disminución de tamaño en la celdilla de los cristales con el nucleótido (**Tabla 4.7**). El cálculo de los valores para el coeficiente de Matthews es de  $2.34 \text{ \AA}^3 \text{ Da}^{-1}$  para los

cristales de la apoenzima y  $2.24 \text{ \AA}^3 \text{ Da}^{-1}$  para los cristales de los complejos binarios. Considerando en ambos casos la presencia de una sola molécula en la unidad asimétrica, el contenido de solvente sería de 47.4% y 45.0% respectivamente.



**Figura 4.29. Patrones de difracción de IP<sub>5</sub> 2-K<sub>W129A</sub>.** a) Patrón procedente del cristal de IP<sub>5</sub> 2-K<sub>W129A</sub>. b) Patrón de difracción procedente del cristal de IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP.

**Tabla 4.7. Procesado de datos de cristales de IP<sub>5</sub> 2-K<sub>W129A</sub> y de complejos IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP e IP<sub>5</sub> 2-K<sub>W129A</sub>/ADP.** Los valores entre paréntesis indican los valores correspondientes a la capa de mayor resolución.

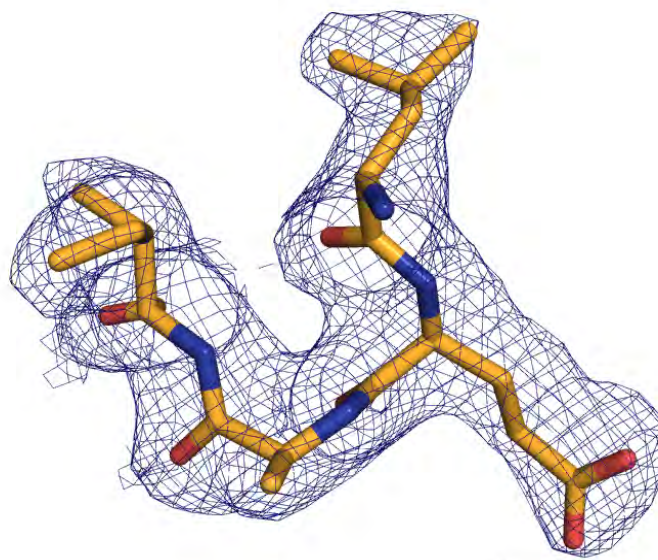
	IP <sub>5</sub> 2-K <sub>W129A</sub>	IP <sub>5</sub> 2-K <sub>W129A</sub> / AMPPNP	IP <sub>5</sub> 2-K <sub>W129A</sub> / ADP
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2		
Moléculas por u.a.	1		
Celdilla (Å)	66.00	63.06	63.72
	68.23	71.80	71.92
	105.80	100.23	100.62
Línea de luz	ID14-4	ID23-2	ID14-4
Longitud de onda (Å)	0.9794	1.0053	0.9334
Resolución (Å)	47.44-2.25	53.37-2.05	39.49-2.50
R <sub>merge</sub> (%)	7.2 (51.7)	9.8 (39.8)	10.5 (40.7)
Reflexiones únicas	24181	29280	16575
Redundancia (%)	7.3 (7.4)	7.1 (7.3)	4.6 (4.8)
Compleitud (%)	99.6 (100)	100 (100)	99.8 (100)
$\langle I/\sigma(I) \rangle$	15.4 (3.8)	15.0 (4.9)	10.1 (3.6)
Wilson B (Å <sup>2</sup> )	47.4	22.0	39.7

$$R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$$

**Tabla 4.8. Reemplazamiento Molecular de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ ,  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /AMPPNP e  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /ADP.** El *TF Z-score* corresponde a la probabilidad de que la solución obtenida sea la solución correcta. Valores  $> 8$  indican la obtención de una solución correcta.

	$\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$	$\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ / AMPPNP	$\text{IP}_5$ 2- $\text{K}_{\text{W129A}}$ / ADP
<i>TF Z-score</i>	23.39	32.57	33.82

Pese a que la homología con la molécula a resolver era prácticamente total (con la única diferencia de la mutación W129A) los primeros intentos de resolver la proteína mediante Reemplazamiento Molecular fueron infructuosos. Dada la posible existencia de un cambio conformacional, se generaron dos “modelos parciales” dividiendo el modelo de búsqueda en dos partes, las cuales correspondían a los lóbulos  $\text{N}_i$  (residuos 3-150) y  $\text{C}_i$  de la proteína (residuos 160-437). Empleando esta aproximación se encontró una única solución de Reemplazamiento Molecular para cada complejo (**Tabla 4.8**). Esta aproximación se llevó a cabo tanto para resolver la apoenzima como para resolver los complejos  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /AMPPNP e  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ /ADP. Los mapas obtenidos mostraron densidades claramente interpretables en todos los casos (**Figura 4.30**).



**Figura 4.30. Sección a  $1\sigma$  del mapa de densidad electrónica de  $\text{IP}_5$  2- $\text{K}_{\text{W129A}}$ .**

Una vez determinadas las diferentes estructuras fueron refinadas. Las estadísticas para el refinamiento pueden observarse en la **Tabla 4.9**.

**Tabla 4.9.** Estadísticas del refinamiento de IP<sub>5</sub> 2-K<sub>W129A</sub> y los complejos IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPNP e IP<sub>5</sub> 2-K<sub>W129A</sub>/ADP.

	IP <sub>5</sub> 2-K <sub>W129A</sub>	IP <sub>5</sub> 2-K <sub>W129A</sub> / AMPNP	IP <sub>5</sub> 2-K <sub>W129A</sub> / ADP
<b>Refinamiento</b>			
Resolución (Å)	47.44-2.25	53.37-2.05	39.49-2.50
R <sub>work</sub> (%)	24.0	20.0	22.0
R <sub>free</sub> (%)	28.0	25.0	28.0
<b>Número de átomos (no H)</b>			
Totales	3246	3839	3515
Proteína	3151	3458	3391
Ligando	0	31	27
Iones	21	26	31
Moléculas de agua	74	324	66
<b>Desviación cuadrática media (rmsd)</b>			
Enlaces (Å)	0.009	0.008	0.005
Ángulos (°)	1.69	1.69	1.04
<b>Factor térmico B (Å<sup>2</sup>)</b>			
Total	56.90	24.90	41.08
Proteína	59.27	24.15	41.10
Ligandos	-	19.10	32.15
<b>Diagrama de Ramachandran</b>			
Residuos en la zona más favorable (%)	92.5	91.2	92.3
<b>Residuos no modelados</b>			
	1:2	386	158:159
	47:60	434:451	334:336
	152:161		386
	334:341		434:451
	382:287		
	434:451		

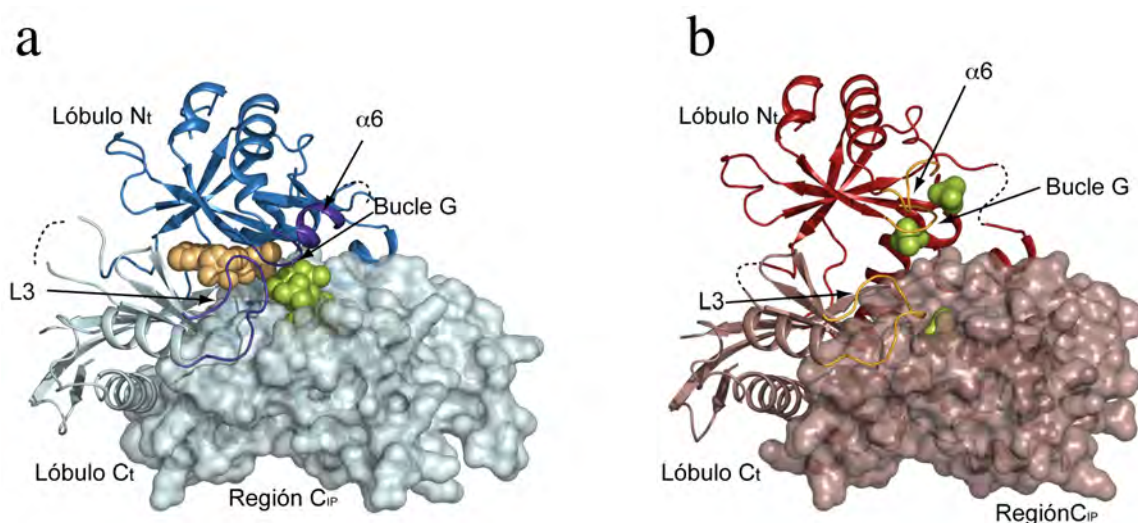
$$R_{\text{work}}/R_{\text{free}} = \sum_{hkl} |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{scale}}| / \sum_{hkl} F_{hkl}^{\text{obs}}$$

## 4.2.5 Estructura de la IP<sub>5</sub> 2-K<sub>W129A</sub>

### 4.2.5.1 Estructura de la IP<sub>5</sub> 2-K<sub>W129A</sub>: conformación abierta

La estructura de la IP<sub>5</sub> 2-K<sub>W129A</sub> presenta un plegamiento general similar al mostrado por la IP<sub>5</sub> 2-K, si bien se observan importantes diferencias (**Figura 4.31**).



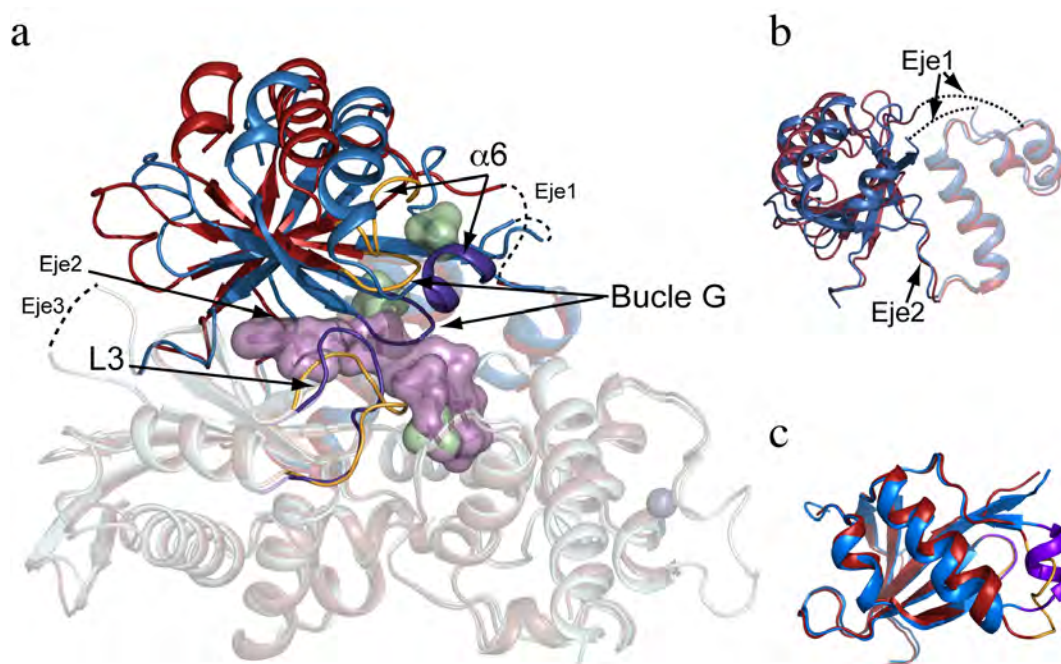


**Figura 4.31. Comparación de las estructuras de la IP<sub>5</sub> 2-K en presencia y ausencia de inositido.** a) Estructura de IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP. El código de colores es el explicado para la **Figura 4.26**. El lóbulo C<sub>IP</sub> se encuentra representado en forma de superficie. b) Estructura de la IP<sub>5</sub> 2-K-W<sub>129A</sub>. El lóbulo N<sub>t</sub> se representa mediante cintas de color rojo, el lóbulo C<sub>t</sub> mediante cintas de color salmón, mostrándose el lóbulo C<sub>IP</sub> como superficie. Se destacan en color naranja los elementos α6, L3 y bucle G. Las superficies de color verde corresponden a iones sulfato.

La principal diferencia entre ambas formas consiste en que los lóbulos N<sub>t</sub> y C<sub>t</sub> se encuentran significativamente más separados entre sí (**Figuras 4.31 y 4.32a**). La superposición de la cadena principal (Cα) de ambas formas presenta un valor de la desviación cuadrática media (rmsd) de 2.5 Å. Pese a que este valor no es muy elevado puede apreciarse un movimiento significativo entre dos regiones de la proteína (**Figura 4.32a**). La primera de estas regiones está formada por la mayor parte del lóbulo N<sub>t</sub>, (residuos 6-42 y 103-152), mientras que la segunda región abarca todo el lóbulo C<sub>t</sub> (161-432) y una pequeña región del N<sub>t</sub> (residuos 63-102). El movimiento en conjunto puede describirse como una rotación de 18.4° de una de las regiones respecto de la otra. Los puntos de bisagra a través de los cuales se produce el movimiento de dominios son los segmentos 42-63, 102-103 y el propio conector entre ambos lóbulos (153-160) (**Figura 4.32a y b**). De ahora en adelante se denominará a esta estructura forma abierta o forma apo, ya que representa la conformación abierta que muestra la forma libre de la enzima, en contraste con la forma unida a inositido, que presenta una forma cerrada.

Pueden apreciarse también diferencias locales, en concreto en los elementos que conforman la cremallera del centro activo. Así pues, la región de α6 observada se encuentra desestructurada en la forma apo y alejada del C<sub>t</sub>, contrastando con la hélice α6 anteriormente en los complejos que unen inositido (**Figura 4.32a y c**). A su vez el bucle L3 se encuentra en una conformación distinta, dejando el sitio de unión a ADP al descubierto. El bucle G, también se encuentra más alejado del C<sub>t</sub>.

La estructura de la apoenzima presenta tres iones sulfato, localizados en posiciones equivalentes a las que ocupan el  $P\gamma$  del ATP y los fosfatos P1 y P6 en la holoenzima. Estos iones sulfato provienen del tampón empleado durante la cristalización de la enzima (**Figura 4.31b** y **Figura 4.32a**).

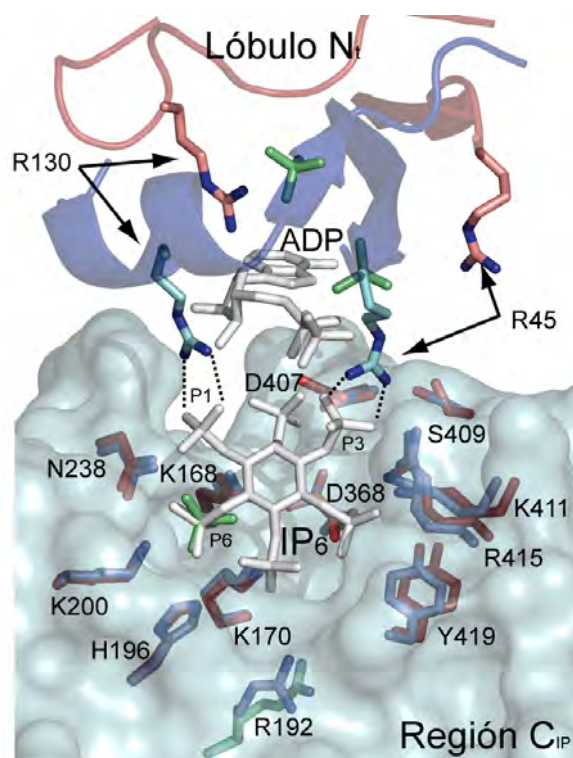


**Figura 4.32. Superposición de las estructuras de IP<sub>5</sub> 2-K en ausencia/presencia de inositido.** a) Representación mediante diagrama de cintas del complejo ternario de IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP (tonos rojos) e IP<sub>5</sub> 2-K<sub>W129A</sub> (tonos azules). Se ha superpuesto el lóbulo C<sub>l</sub> de ambas estructuras (zona transparente). El IP<sub>6</sub> y ADP se representan con superficie morada, y los iones sulfato pertenecientes a la forma apo con superficie verde. b) Comparación de los N<sub>t</sub> según la superposición anterior, mostrándose el cambio de conformación en esta zona y los ejes 1 y 2. c) Cambio de conformación en la zona de la hélice  $\alpha 6$ . La forma cerrada presenta la hélice  $\alpha 6$  (violeta) formada, mientras que la forma abierta mostrada por IP<sub>5</sub>-2K<sub>W129A</sub> presenta un bucle (naranja).

#### 4.2.5.2 Sitio de unión al inositido: comparación entre forma abierta y forma cerrada

Comparando las estructuras de la forma abierta y la cerrada, se pueden observar los cambios que se producen en el centro activo tras la unión de inositido. La mayor parte del lóbulo C<sub>ip</sub> se mantiene inalterado tras la unión del inositido, pudiéndose observar que mayoría de los aminoácidos que aporta el C<sub>ip</sub> para coordinar el IP sufren mínimos cambios (**Figura 4.33**). Este hecho indica que el bolsillo de unión a inositido se encuentra preformado en su mayor parte. No obstante este sitio no está totalmente creado en la forma abierta, pues los residuos que aporta el lóbulo N<sub>t</sub> al centro activo (Arg45, Arg130 y residuos del bucle G) se encuentran demasiado alejados como del sustrato.





**Figura 4.33.** El sitio de unión a inositol se encuentra preformado en la forma apo. Superposición del complejo ternario de  $IP_5$  2- $K/IP_6/ADP$  e  $IP_5$  2- $K_{W129A}$  en el lóbulo  $C_t$  representado mediante superficie azul claro. Los lóbulos  $N_t$  se representan mediante diagrama de cintas en azul y naranja respectivamente. Los residuos marcados en tonos azules corresponden a la forma cerrada, los residuos en colores rojizos a la forma abierta. El  $IP_6$  y ADP se representan con color blanco. Los diferentes residuos que conforman el sitio de unión a inositol del  $C_t$  se encuentran en la misma posición, mientras que los residuos que pertenecen al  $N_t$  se encuentran sustancialmente alejados en la forma apo.

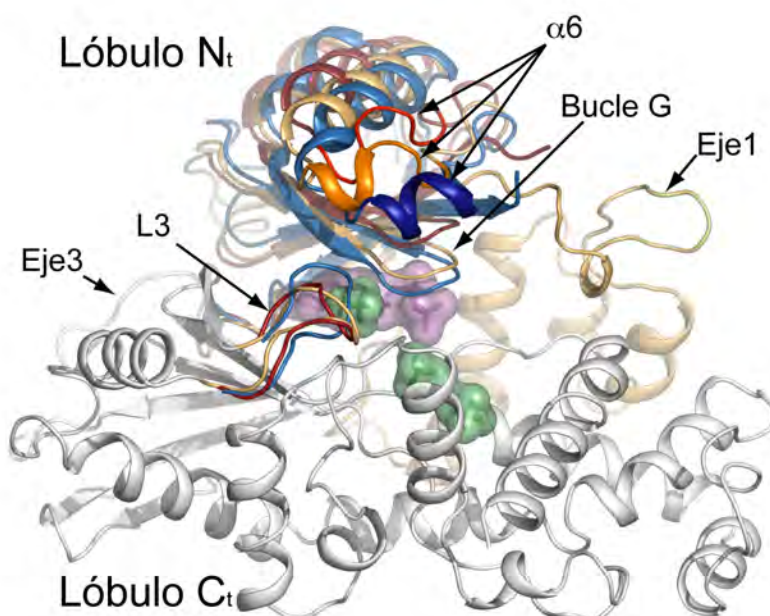
#### 4.2.6 Estructura de los complejos $IP_5$ 2- $K_{W129A}/AMPPNP$ e $IP_5$ 2- $K_{W129A}/ADP$

##### 4.2.6.1 Estructura de la $IP_5$ 2- $K_{W129A}$ unida al nucleótido: conformación semicerrada

Con el fin de esclarecer si el nucleótido también origina algún cambio en la enzima, se cristalizó  $IP_5$  2- $K_{W129A}$  en complejo con AMPPNP y ADP. La enzima unida a AMPPNP o ADP se encuentra en un estado intermedio entre las formas anteriormente descritas, es decir, representa una forma a medio camino entre forma abierta (rmsd =  $1.47^\circ$ ) y forma cerrada (rmsd =  $1.71^\circ$ ) (**Figura 4.34**). La rotación entre dominios es de  $10^\circ$  (comparación entre forma abierta y forma semicerrada). El mecanismo y las regiones que participan en este cambio son las mismas que para el cambio conformacional observado entre la apoenzima y la holoenzima.

Puede observarse que los residuos clave para la coordinación del inositido del  $N_t$  se encuentran más cercanos al lóbulo  $C_t$  que en la forma abierta, si bien aún están apreciablemente lejos para coordinar el inositido. La hélice  $\alpha 6$  encontrada en la forma cerrada, y que se presentaba totalmente desestructurada en la forma abierta (forma apo), se forma parcialmente tras la unión del nucleótido, estableciendo interacciones con L3, si bien la cremallera no está totalmente formada. Puede apreciarse también cómo al producirse la unión del nucleótido, el bucle-G se encuentra más cerca del lóbulo  $C_t$ .

Los cristales de la enzima en conformación semicerrada presentan iones sulfato en el centro activo, en posiciones equivalentes a las que ocuparían los fosfatos P1, P5 y P6 del inositido (**Figura 4.34**).

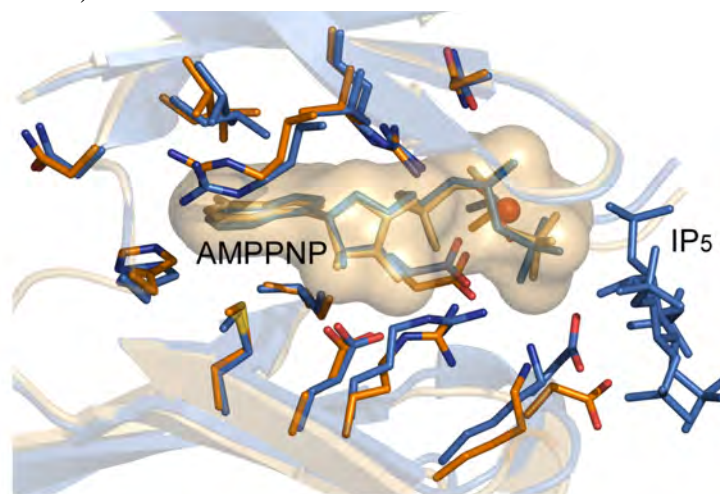


**Figura 4.34. Superposición de las tres formas de  $IP_5$  2-K.** Estructura de  $IP_5$  2- $K_{W129A}$  en complejo con AMPPNP, representada en modelo de cintas, con el lóbulo  $N_t$  en naranja y el  $C_t$  en blanco. La región del  $N_t$  comprendida entre los residuos 63-102 se representa en naranja pálido. Tras superponer los dominios  $C_t$  de cada forma, se muestra en rojo el  $N_t$  de la forma abierta ( $IP_5$  2- $K_{W129A}$ ) y en azul el  $N_t$  de la forma cerrada ( $IP_5$  2-K en complejo con inositido). En superficie violeta se observa el AMPPNP del complejo  $IP_5$  2- $K_{W129A}$ /AMPPNP, y en superficie verde los iones sulfato de ese mismo complejo.

#### 4.2.6.2 Sitio de unión a inositido : comparación entre forma abierta y forma semicerrada

Comparando el sitio de unión a nucleótido entre los diferentes complejos obtenidos ( $IP_5$  2- $K_{W129A}$ /AMPPNP,  $IP_5$  2-K/ $IP_5$ /AMPPNP,  $IP_5$  2-K/ $IP_6$ /ADP e  $IP_5$  2- $K_{W129A}$ /ADP) puede observarse que no hay cambios significativos en éste. No obstante, existen dos diferencias significativas: por un lado, en ausencia del inositido no se observan iones  $Mg^{2+}$  en el centro activo, lo que parece indicar que la presencia del metal se encuentra

asociada a la unión de los dos sustratos, y no exclusivamente a la unión del nucleótido. Por otro lado, existe un ligero reajuste del bolsillo de unión a ATP tras la entrada del inositido (**Figura 4.35**).



**Figura 4.35.** El sitio de unión al nucleótido en ausencia de inositido. Superposición de los complejos IP<sub>5</sub> 2-K<sub>W129A</sub>/AMPPNP e IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP representados como modelo de cintas, en color naranja y azul respectivamente. Los residuos importantes para la coordinación del nucleótido se representan en el color correspondiente. La esfera de color rojo representa al ion Mg<sup>2+</sup> presente en el complejo IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP. Puede apreciarse que los residuos se encuentran en la misma posición en ambos casos.

#### 4.2.7 Formación de la cremallera (interacción $\alpha$ 6-L3) en las diferentes formas de IP<sub>5</sub> 2K

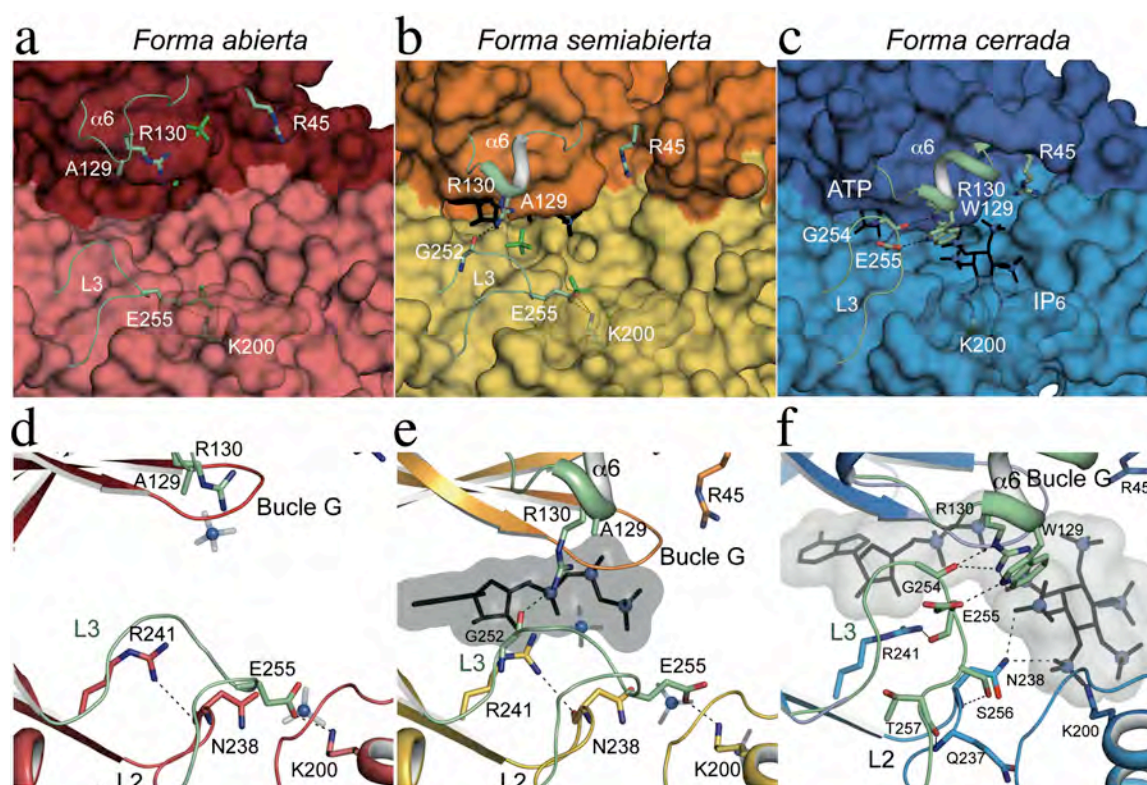
Además del movimiento de apertura y cierre de dominios, IP<sub>5</sub> 2-K experimenta cambios importantes en los elementos que forman la cremallera del centro activo debido a la entrada del nucleótido y el inositido.

Cuando la enzima se encuentra en su forma apo, puede observarse que  $\alpha$ 6 se encuentra desordenada con los residuos que aporta a la cremallera alejados del centro activo. A su vez, L3 se encuentra en una conformación plegada, anclado a la región C<sub>IP</sub> mediante una interacción entre Glu255-Lys200. En esta situación la conformación del bucle L2 se encuentra fijada por la interacción entre Asn238-Arg241 (**Figura 4.36a y d**).

La entrada del nucleótido conlleva cambios principalmente en la zona de  $\alpha$ 6, la cual se forma parcialmente, aproximándose la Arg130 al C<sub>i</sub>. Este acercamiento permite que se forme la primera interacción de la cremallera del centro activo: Arg130-Gly252. El cambio de conformación aproxima también Arg45 al C<sub>i</sub>. En cuanto al bucle G, éste se aproxima también al C<sub>i</sub>, adquiriendo casi su posición definitiva. Paralelamente, tanto el bucle L2 como el L3 siguen manteniendo la misma conformación que en la forma apo (**Figura 4.36b y e**).

La entrada del inositido es la que provoca un mayor número de cambios. El puente salino entre Glu255-Lys200, que mantenía fija la conformación de L3 se rompe. La ruptura de esta interacción libera el Glu255, el cual interacciona en este caso con el Trp129 de  $\alpha 6$ , cerrando la cremallera. En esta cremallera, se forma también una interacción adicional entre Gly254-Arg130, rompiéndose la anteriormente formada entre Gly252-Arg130. Paralelamente, la interacción en L2 formada por Asn238-Arg241, se rompe al entrar el IP. Esto posibilita la aparición de interacciones entre este bucle y L3: Gln237-Thr257 y Asn238-Gln256 (**Figura 4.36c y f**). En conjunto puede observarse que el L3 pasa de una conformación plegada (formas abierta y semiabierta) a una conformación extendida (forma cerrada).

Por tanto, puede decirse que la entrada del inositido conlleva el cierre de esta cremallera.



**Figura 4.36. Cambios en los elementos  $\alpha 6$ /L3.** La estructura de la  $IP_5$  2- $K_{W129A}$  está representada mediante tonos rojos, la estructura de  $IP_5$  2- $K_{W129A}$ /AMPPNP está representada mediante tonos naranjas y la estructura de  $IP_5$  2-K/ $IP_6$ /ADP mediante tonos azules. Los paneles de la zona superior representan las enzimas en forma de superficie y los de la zona inferior mediante un modelo de cintas. En los tres casos, los elementos pertenecientes a  $\alpha 6$ /L3 se representan en color verde. Los sustratos y productos se representan mediante color negro. a y d) Disposición del entorno del centro activo de  $IP_5$  2- $K_{W129A}$  (forma apo o abierta). b y e) Disposición del entorno del centro activo de  $IP_5$  2- $K_{W129A}$ /AMPPNP (forma semicerrada). c y f) Disposición del entorno del centro activo de  $IP_5$  2-K/ $IP_6$ /ADP (forma cerrada).



## 4.2.8 Análisis mediante mutagénesis dirigida y ensayos de actividad

### 4.2.8.1 Análisis mediante mutagénesis dirigida de la cremallera $\alpha$ 6-L3 y residuos del N<sub>t</sub> implicados en la unión a IP

Con el objetivo de confirmar los papeles de residuos identificados en este estudio, se obtuvieron diferentes mutantes de la IP<sub>5</sub> 2-K cuya actividad enzimática se ha medido. Por un lado, se han mutado los residuos implicados en la formación de la cremallera: además del mutante W129A se realizó el mutante R130A, el mutante doble W129V/R130A y la delección de los residuos Ser253/Glu255 ( $\Delta$ S253/ $\Delta$ E255). Por otro lado, y con el fin de analizar el papel del P3 del inositido en la obtención de la forma cerrada de la enzima y/o en la formación de la cremallera, se obtuvo el mutante R45A. La actividad de cada uno de estos mutantes en relación a la proteína nativa puede encontrarse en la **Tabla 4.10**.

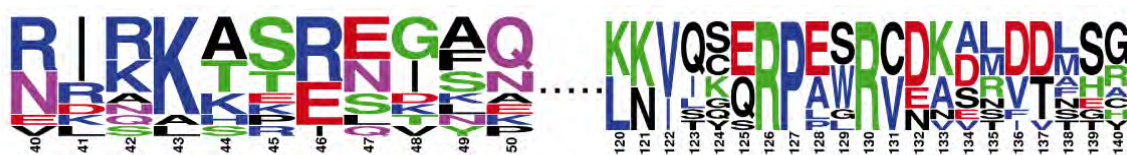
**Tabla 4.10. Actividad enzimática relativa frente a IP<sub>5</sub> de IP<sub>5</sub> 2-K y diferentes mutantes.** Los ensayos se llevaron a cabo empleando una concentración de 4  $\mu$ M de IP<sub>5</sub> y 100  $\mu$ M de ATP. En cada ensayo se realizaron un total de 5 medidas independientes, repitiéndose el ensayo completo en 3 ocasiones.

Muestra	Actividad (%)
Nativa	100
R45A	58 $\pm$ 10
W129A	121 $\pm$ 11
W129V/R130A	21 $\pm$ 11
R130A	16 $\pm$ 9
$\Delta$ S253/ $\Delta$ E255	327 $\pm$ 36

Los ensayos mostraron que la actividad del mutante W129A permanecía inalterada, siendo incluso ligeramente superior a la actividad de la enzima nativa. Esto indica que el Trp129 no es crítico para la catálisis. Por el contrario, la mutación de Arg130 disminuye drásticamente la actividad enzimática. El doble mutante W129V/R130A mostró también una caída importante en su actividad, lo que confirma el papel clave de la Arg130. La función de este residuo es doble: por un lado participa en la formación de la cremallera pudiendo ser esencial en la formación de la forma cerrada y, por otro, coordina el P1 del inositido. Por último la delección  $\Delta$ S253/ $\Delta$ E255 mostró una actividad 3 veces superior a la de la enzima nativa. Estos dos residuos pertenecen al bucle L3, que se encuentra bloqueando el centro activo. Por tanto, la eliminación de estos residuos

suprimiría el bloqueo del centro activo, lo que sería coherente con una salida de productos más rápida y una mejora en la cinética de reacción.

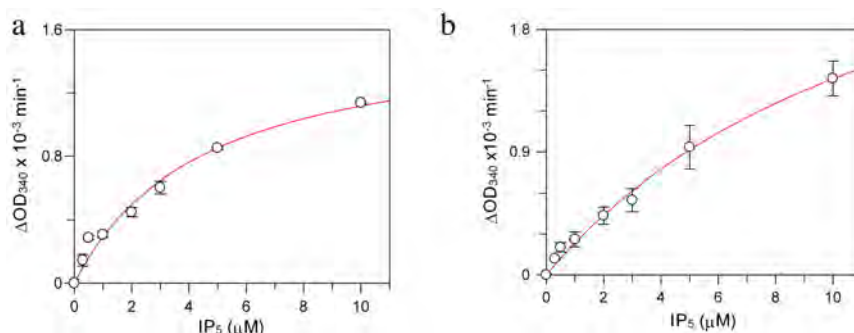
Por último, se analizó la actividad del mutante R45A. La actividad decrece al mutar este residuo, pero IP<sub>5</sub> 2-K conserva una capacidad catalítica significativa frente a IP<sub>5</sub>. Esto demuestra que la coordinación de P3 mediante la Arg45 no es esencial para la catálisis. Este residuo (al contrario que la Arg130) no está conservado en IP<sub>5</sub> 2-Ks de otras especies (**Figura 4.37**).



**Figura 4.37.** Análisis de la secuencia de diferentes especies de IP<sub>5</sub>-2K (*sequence logos*). El tamaño de cada residuo representa su abundancia relativa en el alineamiento. Puede observarse que mientras que la Arg130 está conservada en todas las especies, el Trp129 sólo se encuentra presente en algunas de las secuencias alineadas.

#### 4.2.8.2 Ensayos cinéticos de IP<sub>5</sub> 2-K e IP<sub>5</sub> 2-K<sub>W129</sub>

Como se explicó anteriormente, el Trp129 no es un residuo crítico para la catálisis, conservando el mutante W129A los niveles de actividad de la enzima nativa (**Tabla 4.10**). Con el fin de realizar una caracterización más profunda de dicho mutante, se realizaron ensayos cinéticos según se ha descrito en Materiales y Métodos (Sección 3.2.1.3) (**Figura 4.38**). Estos ensayos revelaron que la enzima nativa presenta una afinidad aparente ( $K_{m\text{ ap}}$ ) tres veces menor por el inositido, mientras que la velocidad máxima aparente ( $V_{\text{max ap}}$ ) es del doble para el mutante que para la enzima nativa (**Tabla 4.11**). Una posible explicación a estos resultados es que la eliminación del Trp129 afecte parcialmente a la formación del bolsillo de unión a inositido, alterándose el reconocimiento de inositido, y, por tanto, la afinidad de la enzima por el sustrato. Sin embargo, la eliminación de esta cadena lateral debilitaría el cierre de la cremallera, facilitando (al igual que ocurría con la eliminación de los residuos Ser253 y Glu255) que la liberación de los productos sea más rápida, lo que hace que aumente la eficiencia de la enzima. En cualquier caso, estas variaciones no son muy significativas y puede decirse que la mutación W129A no afecta a la actividad enzimática produciéndose variaciones pequeñas en los parámetros cinéticos.



**Figura 4.38. Ensayos cinéticos de IP<sub>5</sub> 2-K.** La gráfica representa la variación de absorbancia a 340 nm a diferentes concentraciones de sustrato. a) Ensayo correspondiente a la enzima nativa. b) Ensayo correspondiente al mutante W129A.

**Tabla 4.11. Parámetros cinéticos de IP<sub>5</sub> 2-K e IP<sub>5</sub> 2-K<sub>W129A</sub>.** Los parámetros fueron determinados mediante un ajuste a la ecuación de Michaelis-Menten.

	$V_{\max \text{ ap}}$ ( $\mu\text{mol}/\text{min}/\text{mg}$ )	$K_{\text{m ap}}$ ( $\mu\text{M}$ )
Nativa	$0.54 \pm 0.07$	$3.7 \pm 1.0$
W129A	$0.95 \pm 0.16$	$13.7 \pm 3.0$

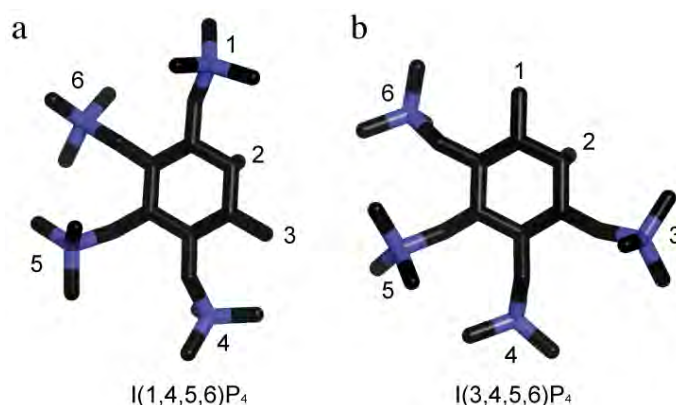
## 4.2.9 Papel del lóbulo N<sub>t</sub> en la catálisis enzimática

El lóbulo N-terminal participa en la unión del inositido y formación de la cremallera, y por tanto parece tener un papel importante en la estabilización de la forma cerrada. En concreto, los elementos que aporta el lóbulo N<sub>t</sub> para la coordinación del IP<sub>5</sub> son, por un lado la Arg130, que coordina el P1 y forma la cremallera, y por otro lado la Arg45 y bucle G, que coordinan P3 (Figura 4.36). Mediante ensayos de mutagénesis, se ha podido comprobar que la Arg130 es esencial para la actividad enzimática mientras que la Arg45, aunque su mutación disminuye moderadamente la actividad enzimática, no lo es. También se ha analizado la unión a IP<sub>5</sub> 2-K de dos isómeros de inositol tetraquisfosfato (IP<sub>4</sub>) que carecen de fosfatos en la posición 1 (I(3,4,5,6)P<sub>4</sub>) o en la posición 3 (I(1,4,5,6)P<sub>4</sub>) con diferentes aproximaciones que se explican a continuación.

### 4.2.9.1 Análisis estructural del complejo IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP

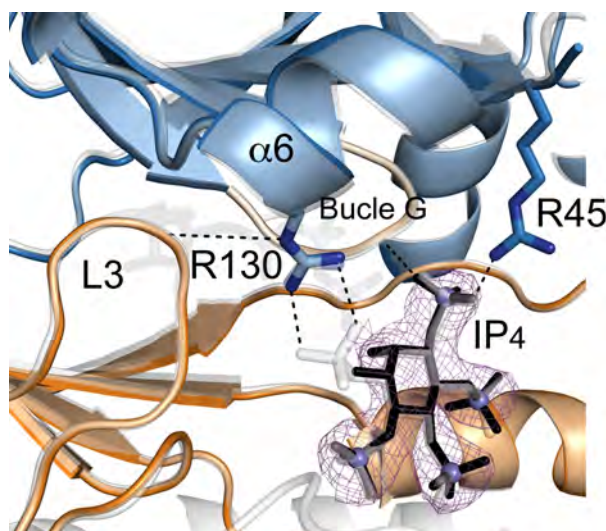
Se realizaron ensayos de cristalización de la enzima con dos isómeros diferentes del IP<sub>4</sub> (I(3,4,5,6)P<sub>4</sub> e I(1,4,5,6)P<sub>4</sub>), que son compuestos enantioméricos (**Figura 4.39**). Se obtuvieron cristales del complejo IP<sub>5</sub> 2-K/I(3,4,5,6)P<sub>4</sub>/AMPPNP y se ha resuelto su estructura (**Figura 4.6** y **Figura 4.40**) (Sección 4.1.3). La estructura obtenida presenta

una conformación cerrada, muy similar a las obtenidas previamente para los restantes complejos ternarios ( $\text{rmsd} = 0.45 \text{ \AA}$ ), encontrándose el inositido en la misma orientación que el  $\text{IP}_5$ . La característica más interesante de este complejo es que la Arg130 mantiene la misma orientación que en los complejos con  $\text{IP}_5$  o  $\text{IP}_6$ , a pesar de la ausencia de P1 de este compuesto. En esta posición, la Arg130 mantiene las interacciones con el L3, de manera que la cremallera de cierre del centro activo está formada del mismo modo que en el resto de complejos con  $\text{IP}_5$  o  $\text{IP}_6$  (**Figura 4.40**). Esto sugiere que la coordinación del P1 no está relacionada con cambio conformacional ni con la formación de la cremallera.



**Figura 4.39. Enantiómeros  $\text{I}(1,4,5,6)\text{P}_4$  e  $\text{I}(3,4,5,6)\text{P}_4$ .** Los fosfatos se encuentran destacados en color azul. a)  $\text{I}(1,4,5,6)\text{P}_4$ . b)  $\text{I}(3,4,5,6)\text{P}_4$ .

No se logró obtener cristales de  $\text{IP}_5$  2-K en complejo con  $\text{I}(1,4,5,6)\text{P}_4$ . Esto podría indicar que la coordinación del P3 sí podría ser necesaria para la obtención de la conformación cerrada de la enzima.



**Figura 4.40. Complejo  $\text{IP}_5$  2-K/ $\text{I}(3,4,5,6)\text{P}_4$ /AMPPNP.** Se ha superpuesto la estructura del complejo ternario con  $\text{IP}_5$  (color blanco) con la estructura de  $\text{IP}_5$  2-K/ $\text{I}(3,4,5,6)\text{P}_4$ /AMPPNP ( $\text{N}_t$  de color azul claro y  $\text{C}_t$  de color naranja claro). Los ligandos se muestran en blanco ( $\text{IP}_5$ ) y negro ( $\text{I}(3,4,5,6)\text{P}_4$ ). La figura muestra también el mapa de densidad a  $1\sigma$  en torno al  $\text{I}(3,4,5,6)\text{P}_4$ .



#### 4.2.9.2 Discriminación de los diferentes enantiómeros de IP<sub>4</sub> por la IP<sub>5</sub> 2-K

Además del IP<sub>5</sub>, la IP<sub>5</sub> 2-K puede emplear como sustratos otros IPs. En relación a esto, se ha observado que los enantiómeros I(1,4,5,6)P<sub>4</sub> e I(3,4,5,6)P<sub>4</sub> son sustratos *in vitro* de diferentes IP<sub>5</sub> 2-K de plantas (Sun *et al.*, 2007). Con el objetivo de profundizar en los determinantes estructurales que rigen el uso de uno u otro compuesto como sustrato, se llevaron a cabo ensayos de actividad frente a ambos sustratos tanto para la enzima nativa como para algunos de los mutantes implicados en la coordinación del P1, P3 y la formación de la cremallera (W129A, R45A y R130A). Los resultados de los ensayos se encuentran recogidos en la **Tabla 4.12**.

**Tabla 4.12. Discriminación de enantiómeros en IP<sub>5</sub> 2-K nativa y diferentes mutantes.**

Muestra	% Uso I(1,4,5,6)P <sub>4</sub>	% Uso I(3,4,5,6)P <sub>4</sub>	Ratio uso I(1,4,5,6)P <sub>4</sub> / I(3,4,5,6)P <sub>4</sub>
Nativa	100	100	0.99±0.15
R45A	179±23	135±8	1.31±0.17
W129A	75±7	97±15	0.77±0.12
R130A	17±7	59±5	0.28±0.07

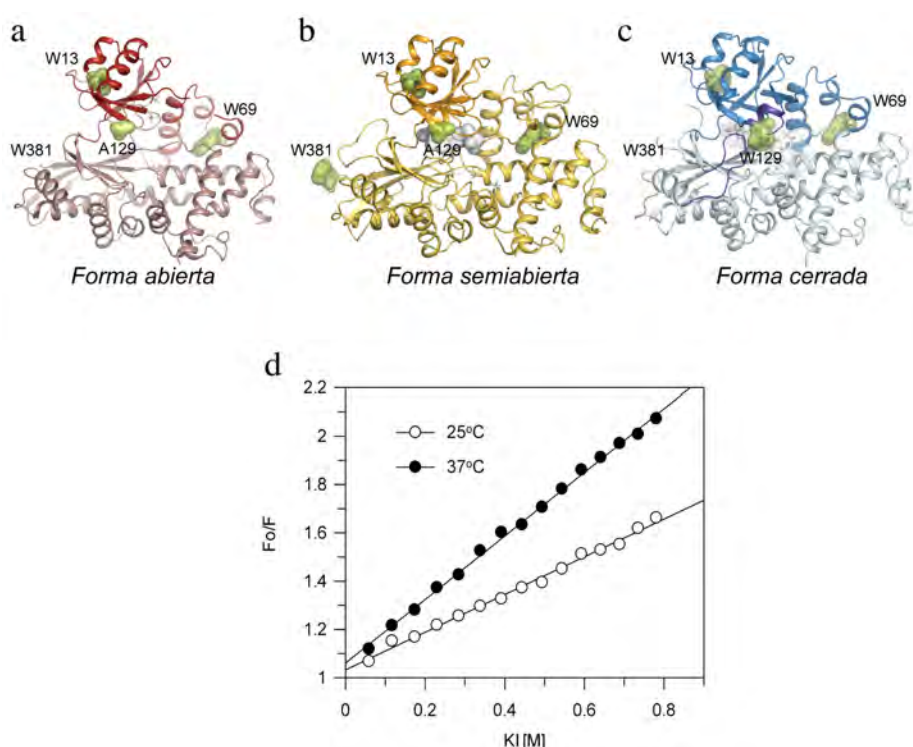
En el caso de *A. thaliana*, la enzima nativa emplea ambos compuestos por igual, mientras que los diferentes mutantes muestran variaciones en la preferencia por uno u otro compuesto. Así, el mutante R45A usa preferentemente I(1,4,5,6)P<sub>4</sub>, mientras que el mutante W129A usa preferentemente I(3,4,5,6)P<sub>4</sub>. Esta tendencia se ve incrementada con la substitución de la Arg130 por alanina, lo que señala de nuevo la importancia de la Arg130 en la coordinación del P1.

El estudio aquí presentado revela por primera vez en la familia de IPKs que la discriminación enantiomérica puede depender de la naturaleza de un único residuo.

## 4.2.10 Otras pruebas experimentales del cambio conformacional

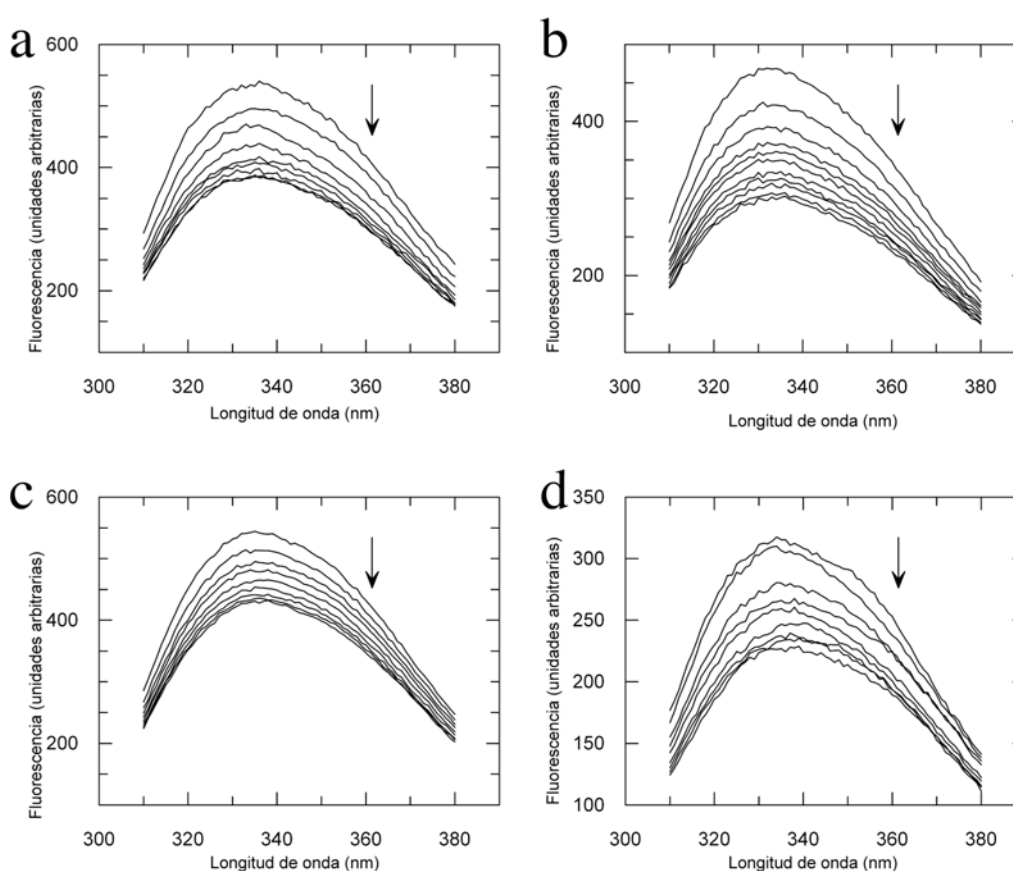
### 4.2.10.1 Fluorescencia intrínseca de la IP<sub>5</sub> 2-K y la IP<sub>5</sub> 2-K<sub>W129A</sub>

La resolución de la estructura de los distintos complejos de IP<sub>5</sub> 2-K ha revelado los cambios de conformación que la enzima sufre durante el ciclo catalítico. Con el fin de estudiar estos cambios se realizaron ensayos de unión a ligando monitorizados mediante espectroscopía de fluorescencia. Para ello se ha analizado la fluorescencia intrínseca de la proteína. La IP<sub>5</sub> 2-K en su forma nativa presenta 4 triptófanos: Trp13 y Trp69 (parcialmente enterrados en bolsillos hidrofóbicos), Trp129 (probablemente cambia su conformación estando expuesto en la apoenzima y parcialmente ocluido en la holoenzima) y Trp381 (totalmente expuesto y desordenado en las formas abierta y semiabierta) (**Figura 4.41a b y c**).



**Figura 4.41.** Ensayos de fluorescencia intrínseca de IP<sub>5</sub> 2-K. a) Posición de los residuos de triptófano en la forma abierta. b) Posición de los residuos de triptófano en la forma semiabierta. c) Posición de los residuos de triptófano en la forma cerrada. d) Análisis de Stern-Volmer para la extinción de los triptófanos en la muestra nativa, empleando ioduro de potasio (KI) como agente de extinción.  $F_0$  es la fluorescencia en ausencia del agente de apagamiento, mientras que  $F$  es la fluorescencia emitida en presencia del agente de apagamiento. La linealidad a diferentes temperaturas indica que los triptófanos se encuentran expuestos al solvente.

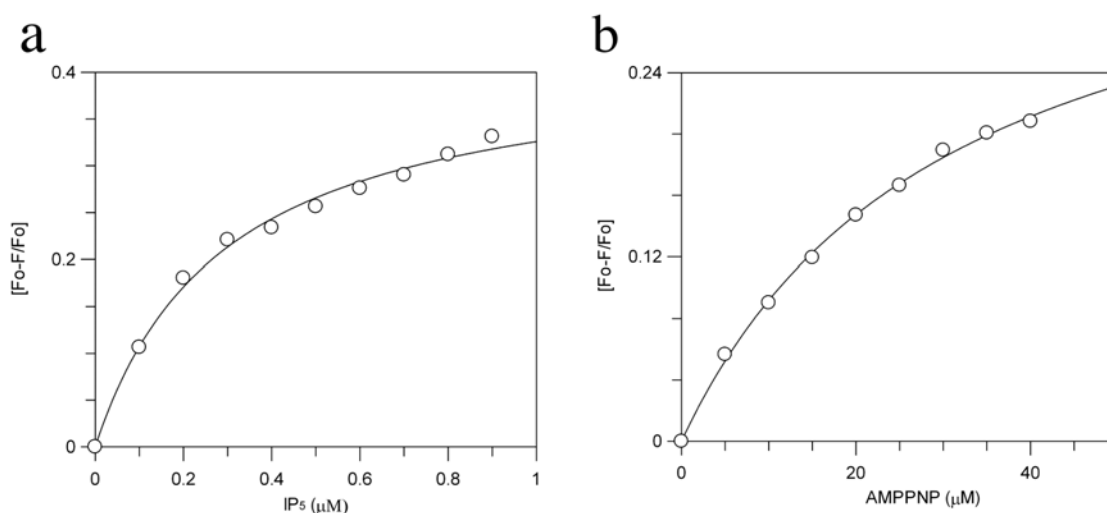
Previamente a los ensayos de unión se determinó que todos los triptófanos (tanto en la enzima nativa como en el mutante W129A) se encontraban expuestos al solvente (**Figura 4.41d**). Los ensayos de unión muestran que para ambas enzimas existe una disminución de la fluorescencia según se realizan adiciones sucesivas de  $IP_5$ , siendo estas disminuciones moderadas en ambos casos (**Figura 4.42a y b**). Asumiendo que el Trp381 no es uno de los principales contribuyentes a estas variaciones (puesto que se encuentra expuesto en todo momento), las variaciones que se observan en la fluorescencia pueden deberse a los ajustes locales en los bolsillos del Trp13 y el Trp69. Estos cambios se producen en ambas muestras. Los mismos resultados se obtienen también al emplear AMPPNP (**Figura 4.42c y d**).



**Figura 4.42. Ensayos de unión a ligando de la  $IP_5$  2-K.** Se representa la fluorescencia (en unidades arbitrarias) frente a la longitud de onda. Las flechas indican la dirección del cambio de la fluorescencia con las adiciones sucesivas de  $IP_5$  o AMPPNP. a) Enzima nativa +  $IP_5$  (0-2.5 mM). b) Mutante W129A +  $IP_5$  (0-2.5 mM). c) Enzima nativa + AMPPNP (0-2.5 mM). d) Mutante W129A + AMPPNP (0-2.5 mM).

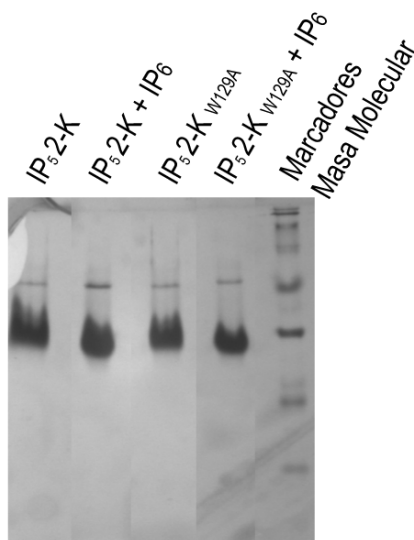
Los experimentos con el sustrato han permitido medir los valores de la constante de disociación ( $K_d$ ). Estos son, en el caso de la proteína nativa,  $0.35 \pm 0.12 \mu M$  (para el  $IP_5$ ) (**Figura 4.43a**) y  $37 \pm 9 \mu M$  (para AMPPNP) (**Figura 4.43b**) y, en el caso del mutante W129A,  $0.42 \pm 0.03 \mu M$  (para el  $IP_5$ ) y  $55 \pm 2 \mu M$  (para AMPPNP). En conjunto estos datos indican que tanto la unión del  $IP_5$  como la unión del ATP provocan cambios

conformacionales en ambos casos, si bien, los cambios en dicha fluorescencia no pueden atribuirse exclusivamente a un posible cambio de conformación del Trp129.



**Figura 4.43. Representación fraccional de los cambios en la emisión de fluorescencia para la proteína nativa.** Valores a una longitud de onda de 335 nm. a) Cambios frente a concentraciones crecientes de  $IP_5$ . b) Cambios frente a concentraciones crecientes de AMPPNP.  $F_0$  es la fluorescencia en ausencia del correspondiente sustrato mientras que  $F$  es la fluorescencia emitida en presencia de dicho sustrato.

#### 4.2.10.2 Ensayos mediante geles nativos



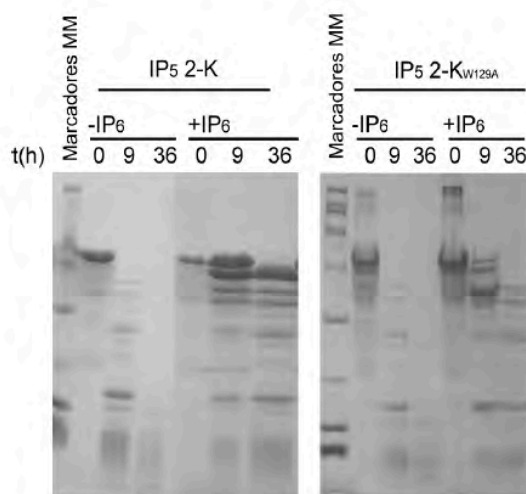
**Figura 4.44. Gel nativo de las diferentes muestras de  $IP_5$  2-K.** Las muestras con  $IP_6$  (tanto de la proteína nativa como de la proteína mutante) muestran una mayor migración en el gel.

Con el fin de corroborar un comportamiento equivalente de la muestra correspondiente al mutante W129A ( $IP_5$  2- $K_{W129A}$ ) e  $IP_5$  2-K nativa se optó por la

realización de geles nativos para muestras de proteína nativa y del mutante W129A, tanto en ausencia como en presencia de IP<sub>6</sub>. Los ensayos realizados muestran un cambio significativo en la migración de las muestras que contienen el inositido. Dicho cambio se produce tanto para la enzima nativa como para la enzima mutante, lo cual confirma que ambas muestras mantienen su capacidad de unión al IP<sub>6</sub>, mostrando ambas un desplazamiento similar en el gel (**Figura 4.44**).

### 4.2.10.3 Ensayos de proteólisis limitada de IP<sub>5</sub> 2-K e IP<sub>5</sub> 2-K<sub>W129A</sub>

Otra técnica también utilizada en la detección de cambios conformacionales en proteínas es la proteólisis limitada (Chu *et al.*, 1997; Jamison *et al.*, 1998; Mueller-Planitz *et al.*, 2013). Se han realizado experimentos de proteólisis limitada con tripsina de la muestra nativa y mutante, analizado su patrón de digestión mediante SDS-PAGE (**Figura 4.45**). El patrón de bandas obtenido difiere entre las muestras que contienen IP<sub>6</sub> y aquellas que no lo contienen, mostrando las muestras que lo contienen una mayor resistencia ante la proteólisis. Esta mayor resistencia a la proteólisis podría deberse a la menor susceptibilidad frente a las proteasas de la forma cerrada (más compacta y con menor superficie expuesta). Este resultado se obtiene tanto para la enzima nativa como para el mutante, lo cual corrobora la existencia de un cambio conformacional inducido por la entrada del inositido en el sitio activo. Resultados similares han sido obtenidos por otros grupos (Gosein *et al.*, 2012).



**Figura 4.45. SDS-PAGE de los ensayos de proteólisis limitada de IP<sub>5</sub> 2-K e IP<sub>5</sub> 2-K<sub>W129A</sub>.** Los ensayos fueron realizados para ambas muestras tanto en presencia como en ausencia del inositido, tomando muestras a diferentes tiempos (0, 9 y 36 horas). La digestión de las muestras fue realizada con tripsina. Tanto en el caso de la proteína nativa como en el de la proteína mutante la adición de IP<sub>6</sub> confiere resistencia a la proteólisis.

### 4.2.11 El ciclo catalítico de la IP<sub>5</sub> 2-K

Uno de los aspectos más destacables de los resultados aquí expuestos es el hecho de que el inositido es un elemento necesario y suficiente para provocar el cambio de conformación de la forma abierta a la forma cerrada, pudiendo ocurrir en ausencia del nucleótido. Puesto que en la forma cerrada el inositido no permite la entrada del nucleótido, éste podría reflejar la capacidad de inhibición por sustrato (IP<sub>5</sub>) o producto (IP<sub>6</sub>), de forma que en condiciones de altos niveles de estos compuestos la enzima se encontraría en conformación cerrada, no siendo posible la catálisis. Fisiológicamente, este proceso de inhibición se daría en situaciones con bajos niveles de ATP y/o altos niveles de IP<sub>6</sub> teniendo como objetivo último el “ahorro energético” a nivel intracelular. Existen múltiples ejemplos de inositol kinasas que presentan este tipo de inhibición por el inositido (Abdullah *et al.*, 1992; Voglmaier *et al.*, 1996; Nalaskowski *et al.*, 2003), pues la síntesis de inositoles fosforilados es un proceso energéticamente muy costoso que debe estar estrictamente regulado.

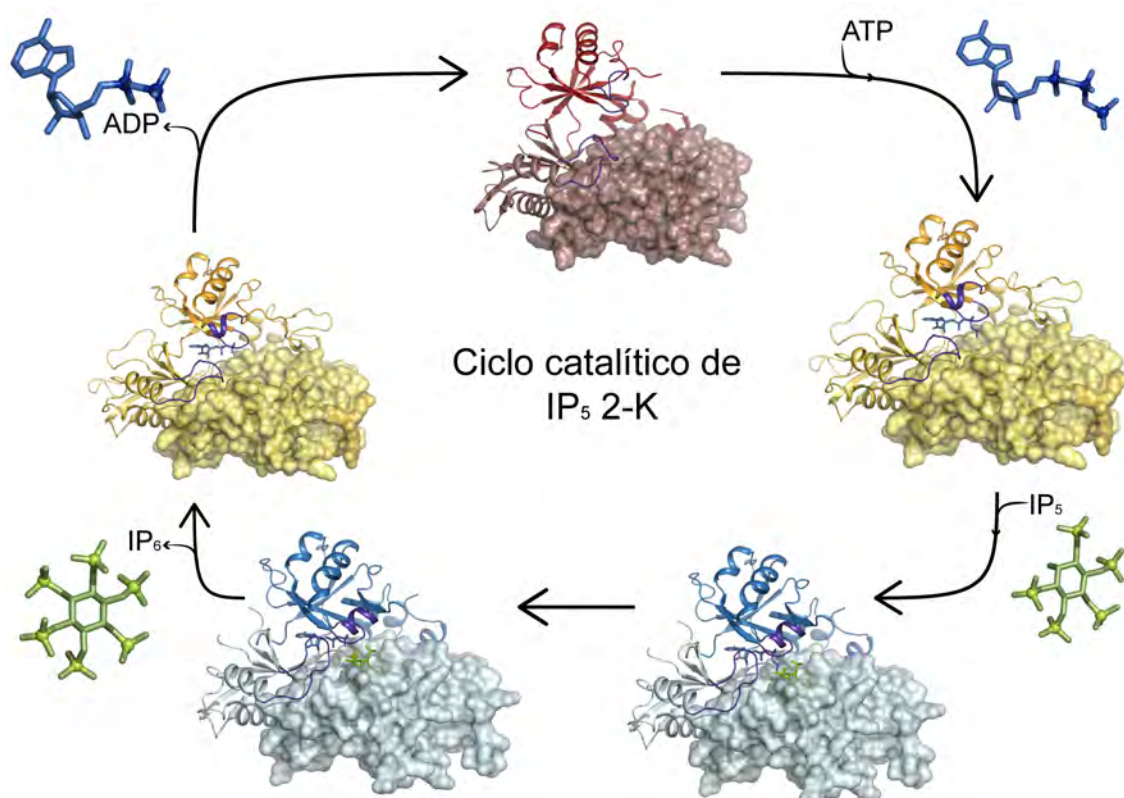
Otro de los efectos del inositido es el cierre de la “cremallera” que bloquea el centro activo. Dicho bloqueo sólo se completa en presencia del inositido. No obstante, la mutación de los diferentes elementos de la cremallera (especialmente aquellos del bucle L3) no afectan a la actividad enzimática, lo que indica que la formación de esta cremallera no es una condición para que se produzca la catálisis. Como se indicó previamente, es probable que este mecanismo de “cremallera” difiera de unas enzimas a otras, ya que los residuos que la forman varían mucho entre especies. Por ejemplo, los residuos de  $\alpha 6$  (y en concreto Trp129 (**Figura 4.37**)) no están conservados. Por otro lado puede observarse que la longitud del bucle L3 varía entre enzimas de diferentes especies, siendo posible que el bucle L3 sea el responsable de las diferencias cinéticas que presentan las enzimas de diferentes especies.

En resumen, tanto el cierre de la cremallera como el cierre completo de los dominios N<sub>i</sub> y C<sub>i</sub> están unidos y supeditados a la entrada de inositido en el centro activo, si bien el cierre de la cremallera no es necesario para que se produzca la actividad.

Las estructuras descritas han permitido elucidar, en base a información estructural, el ciclo catalítico de esta familia de enzimas. Los datos obtenidos sugieren que probablemente el nucleótido se une primero, cerrando parcialmente la estructura y acercando los dominios N<sub>i</sub> y C<sub>i</sub>. Posteriormente entraría el inositido, el cual promovería el cierre completo de ambos dominios y la formación de la cremallera. Esta forma cerrada (holoenzima) revela que el inositido bloquea la entrada y salida del nucleótido del sitio activo, por lo que tras producirse la transferencia de fosfato, el proceso de salida de productos debería ser inverso, abandonando el sitio activo el IP<sub>6</sub>, y posteriormente el ADP (**Figura 4.46**). Tal y como ha sido sugerido para algunas



proteínas kinasas, este último paso de liberación de productos podría ser el paso limitante de la reacción (Cox *et al.*, 1994), pues es necesaria la apertura de ambos dominios y/o la desaparición de la cremallera que cierra el sitio activo. Esta idea se ve reforzada por el hecho de que en dos de los mutantes de diferentes elementos de esta cremallera (W129A y  $\Delta$ S253/ $\Delta$ E255) la eficiencia enzimática es mayor.



**Figura 4.46. Posible ciclo catalítico de la IP<sub>5</sub> 2-K.** El ciclo se propone en base a las estructuras y datos bioquímicos obtenidos durante el desarrollo de este trabajo. La estructura se representa mediante un modelo de cintas, mostrándose la región C<sub>IP</sub> mediante superficie. El color rojo se emplea para la forma apo, el color amarillo para la forma unida a nucleótido (AMPPNP y ADP) y el color azul para los complejos ternarios (IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP e IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP).  $\alpha$ 6 y L3 se representan en color violeta en todos los casos. El IP<sub>5</sub>/IP<sub>6</sub> se representa en verde claro con los fosfatos destacados mediante esferas, mientras que el AMPPNP/ADP se representa en color azul con los fosfatos destacados mediante esferas.

#### 4.2.12 Cambios conformacionales dentro de la familia de las IPKs

Las características estructurales de las diferentes inositol kinasas han sido ampliamente estudiadas con el objetivo de determinar los elementos claves de la especificidad por los sustratos. Múltiples trabajos han mostrado la estructura de diferentes enzimas tanto en su forma libre como unidas a sustratos y nucleótidos. En lo que a la familia de las IPKs se refiere, ninguna de las estructuras resueltas (IP<sub>3</sub> 3-K e IPMK) muestran cambios conformacionales al unirse el nucleótido (González *et al.*,



2004; Miller y Hurley, 2004; Holmes y Jogl, 2006). En el caso de la IP<sub>3</sub> 3-K se observa un ligero cambio si se comparan las estructuras de la enzima en su forma libre y en su forma unida a inositido (González *et al.*, 2004). Este cambio consiste en una reordenación del sitio de unión a inositido, para lo cual un puente salino (Arg419-Asp374) se rompe, lo que permite al residuo Arg419 interactuar con el inositido. En el caso de la IP<sub>5</sub> 2-K, un puente salino (Lys200-Asp255) también se rompe durante la transición entre forma abierta y forma cerrada, aunque no es equivalente al observado en IP<sub>3</sub> 3-K. Hasta la fecha, ni en las diferentes IPKs ni en ninguna otra inositol kinasa han sido caracterizados movimientos de dominios como los observados para IP<sub>5</sub> 2-K. No obstante, aún no se dispone de las estructuras de todas las inositol kinasas (ni tan siquiera de todos los miembros de la familia de las IPKs) en sus diferentes estados, por lo que es posible que otras inositol kinasas presenten cambios conformacionales como los observados para la IP<sub>5</sub> 2-K.

### 4.2.13 Conclusión

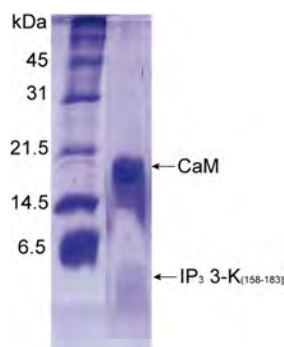
Se ha obtenido la estructura de la forma libre de IP<sub>5</sub> 2-K. Ésta revela que la enzima se encuentra en una conformación más abierta en ausencia de sustratos o productos. Por otro lado, la cristalización de la enzima en presencia exclusivamente del nucleótido (AMPPNP o ADP) ha revelado que la conformación de la enzima en presencia de estos compuestos es un estado intermedio entre la forma abierta (enzima libre) y la forma cerrada (enzima en presencia de inositido). En conjunto, los datos sugieren que IP<sub>5</sub> 2-K experimenta un proceso de apertura y cierre durante las diferentes fases de su ciclo catalítico. Este cambio de conformación lleva asociado cambios en el centro activo, especialmente en los residuos que participan en la interacción entre la  $\alpha 6$  (Leu128-Ala133) y el bucle L3. También se ha caracterizado la actividad de diferentes mutantes de esta zona y se ha estudiado la discriminación de los enantiómeros I(1,4,5,6)P<sub>4</sub> e I(3,4,5,6)P<sub>4</sub> en IP<sub>5</sub> 2-K nativa y diferentes mutantes.

## 4.3 Regulación de la IP<sub>3</sub> 3-K mediante CaM

### 4.3.1 Obtención de muestra de complejos CaM/CBD IP<sub>3</sub> 3-K

Con el fin de cristalizar el complejo entre la calmodulina (CaM) y el dominio de unión a CaM de la IP<sub>3</sub> 3-K (CBD, del inglés “*CaM Binding domain*”) se realizaron tres construcciones diferentes de dicho CBD, entre los aminoácidos 158-201 (IP<sub>3</sub> 3-K<sub>(158-201)</sub>), 158-189 (IP<sub>3</sub> 3-K<sub>(158-189)</sub>) y 158-183 (IP<sub>3</sub> 3-K<sub>(158-183)</sub>). Los niveles de expresión de todos estos fragmentos fueron equivalentes, obteniendo un rendimiento de entre 10-12 mg/L de proteína pura, tras la purificación mediante una columna de *Sefarosa* CL-4B. Igualmente, la CaM se obtuvo con buen rendimiento (20 mg/L de cultivo), eluyendo la proteína de la columna de fenil-Sefarosa cuando el gradiente se encontraba formado por 30% de tampón A (0.6 mM de EGTA). El rendimiento obtenido fue menor que el descrito en el protocolo ya publicado (Hayashi *et al.*, 1998). Esto puede deberse al diferente origen de la proteína expresada (humano en nuestro caso, de rata en el caso del protocolo ya publicado) o también a diferencias en el proceso de la expresión de la proteína.

Con el fin de obtener el complejo, se mezclaron ambas muestras en proporción molar 1:3 (CaM:CBDs). Con este método se pudo aislar el complejo entre ambas proteínas (**Figura 4.47**) mediante exclusión molecular.

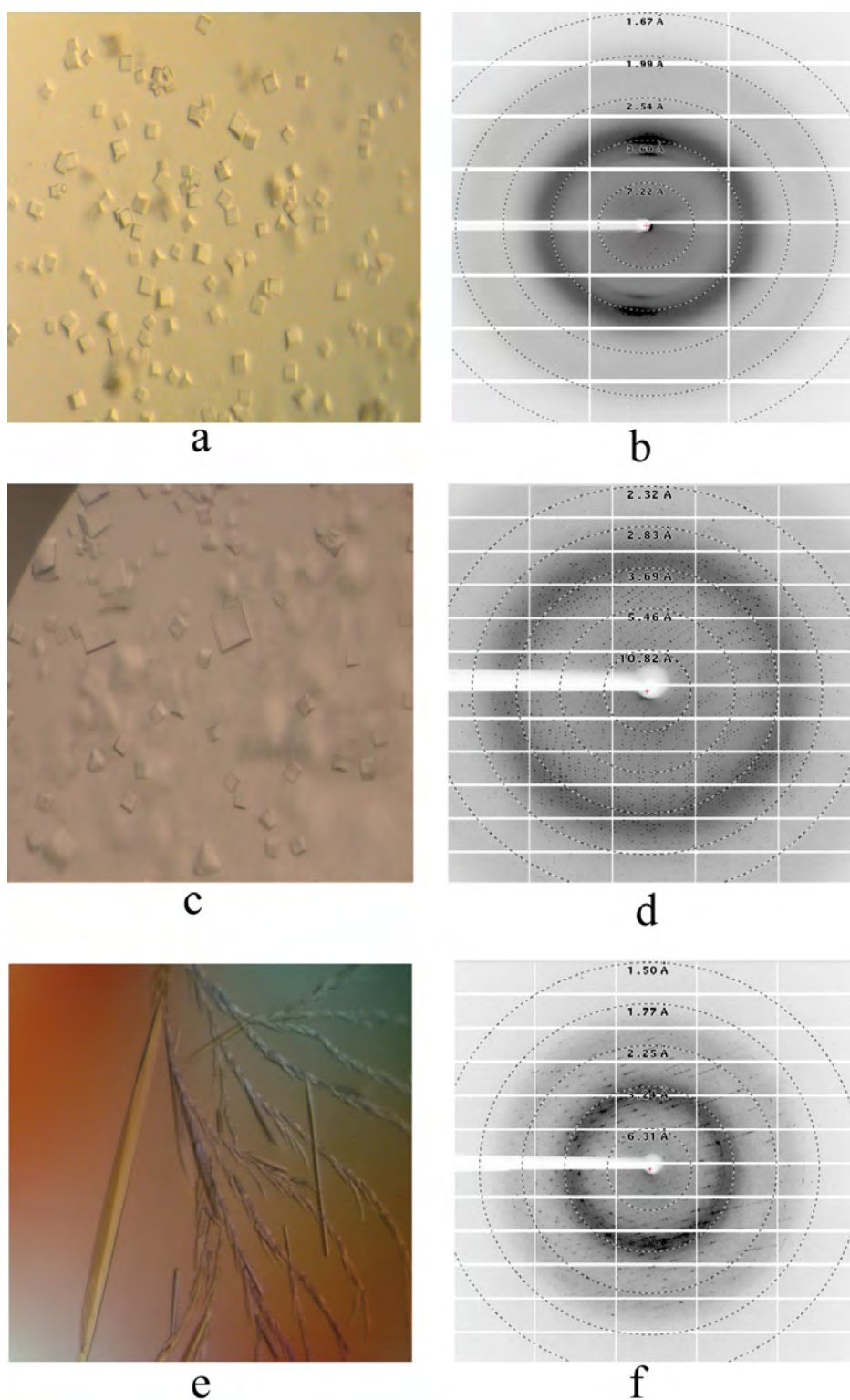


**Figura 4.47.** SDS-PAGE mostrando la muestra pura del complejo IP<sub>3</sub> 3-K<sub>(158-183)</sub>/CaM.

Se produjeron tres tipos de muestras de complejo: complejo CaM/IP<sub>3</sub> 3-K<sub>(158-201)</sub>, complejo CaM/IP<sub>3</sub> 3-K<sub>(158-189)</sub> y complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.

### 4.3.2 Cristalización del complejo CaM/ IP<sub>3</sub> 3-K<sub>(158-183)</sub>

Se realizaron ensayos de cristalización con las tres muestras anteriores, sin embargo únicamente se logró obtener cristales con la muestra del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.



**Figura 4.48. Cristalización del complejo CaM/ IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) Cristales obtenidos con 0.2 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M acetato de sodio pH 4.0 y 29% (m/v) de PEG 4000. b) Patrón obtenido a partir de los cristales mostrados en a). c) Cristales obtenidos en la misma condición y 10 mM de espermidina. d) Patrón obtenido a partir de los cristales mostrados en c). e) Cristales obtenidos en la misma condición y 3.21 mM óxido de n-N,N-dimetilamina. f) Patrón de difracción correspondiente a los cristales mostrados en e).

Se obtuvieron cristales con hábitos distintos en función de las condiciones y aditivos o detergentes utilizados en la cristalización (**Figura 4.48a, c y e**).

### 4.3.3 Recogida de datos, resolución estructural y refinado del modelo

Tras el proceso inicial de optimización los cristales obtenidos mostraron un tamaño adecuado para su estudio por difracción de rayos X (**Figura 4.48a**). No obstante, estos cristales presentaban un patrón de difracción de baja calidad (3.8 Å) (**Figura 4.48b**), aunque suficiente para un análisis inicial (**Tabla 4.13**). Los cristales presentaban un contenido de solvente del 71.24% (considerando la presencia de una molécula en la unidad asimétrica). La estructura fue resuelta mediante Reemplazamiento Molecular (**Tabla 4.14**). El mapa de densidad electrónica obtenido corroboró la presencia de las dos proteínas en el complejo, pero dada la baja resolución no permitió trazar un modelo del CBD. No obstante, la densidad electrónica reveló que este se disponía en forma de hélice  $\alpha$  una vez unido a CaM (**Figura 4.49a**).

**Tabla 4.13** Procesado de datos de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> y CaM.

	Hábito 1 (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 1 + espermidina (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 2 (CaM)
Grupo espacial	I432		P3 <sub>2</sub> 21
Moléculas por u.a.	1		1
Celdilla (Å)	$a=b=c=160.12$	$a=b=c=159.25$	$a=b=40.27$ $c=175.31$
Línea de luz	PXIII (SLS)	XALOC (ALBA)	XALOC (ALBA)
Longitud de onda (Å)	1.0000	0.9794	0.9794
Resolución (Å)	46.18-3.80	79.62-2.34	43.83-1.84
R <sub>merge</sub> (%)	26.7 (64.3)	14.8 (41.2)	8.4 (47.4)
Reflexiones únicas	3693	14923	15231
Redundancia (%)	18.9 (20.0)	17.7 (16.2)	8.0 (8.3)
Compleitud (%)	99.9 (100)	100 (100)	100 (100)
$\langle I/\sigma(I) \rangle$	13.3 (5.9)	12.4 (6.6)	12.6 (4.4)
Wilson B (Å <sup>2</sup> )	32.3	23.4	25.1

$$R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$$

**Tabla 4.14. Reemplazamiento Molecular de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> y CaM.** El *TF Z-score* corresponde a la probabilidad de que la solución obtenida sea la solución correcta. Valores > 8 indican la obtención de una solución correcta

	Hábito 1 (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 1 + espermidina (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 2 (CaM)
<i>Valor Q</i> ( <i>Balbes</i> )	0.77	-----	-----
<i>TF Z-score</i> ( <i>Phaser</i> )	-----	18.19	30.42

Los cristales obtenidos con espermidina mostraron una gran mejoría en cuanto a los límites de difracción (2.34 Å) (**Figura 4.48c y d**). El grupo espacial resultó ser I432, al igual que en los cristales anteriores, con una celdilla de  $a = b = c = 159.25$  Å. Los datos de difracción fueron procesados (**Tabla 4.13**) y la estructura determinada por Reemplazamiento Molecular (**Tabla 4.14**). Los mapas de densidad electrónica obtenidos mostraron buena calidad (**Figura 4.49b**), pudiéndose construir un modelo del complejo casi en su totalidad. Los valores obtenidos tras el refinamiento del modelo se recogen en la **Tabla 4.15**. La zona de mayor desorden se encuentra en los extremos de IP<sub>3</sub> 3-K<sub>(158-183)</sub> y CaM.

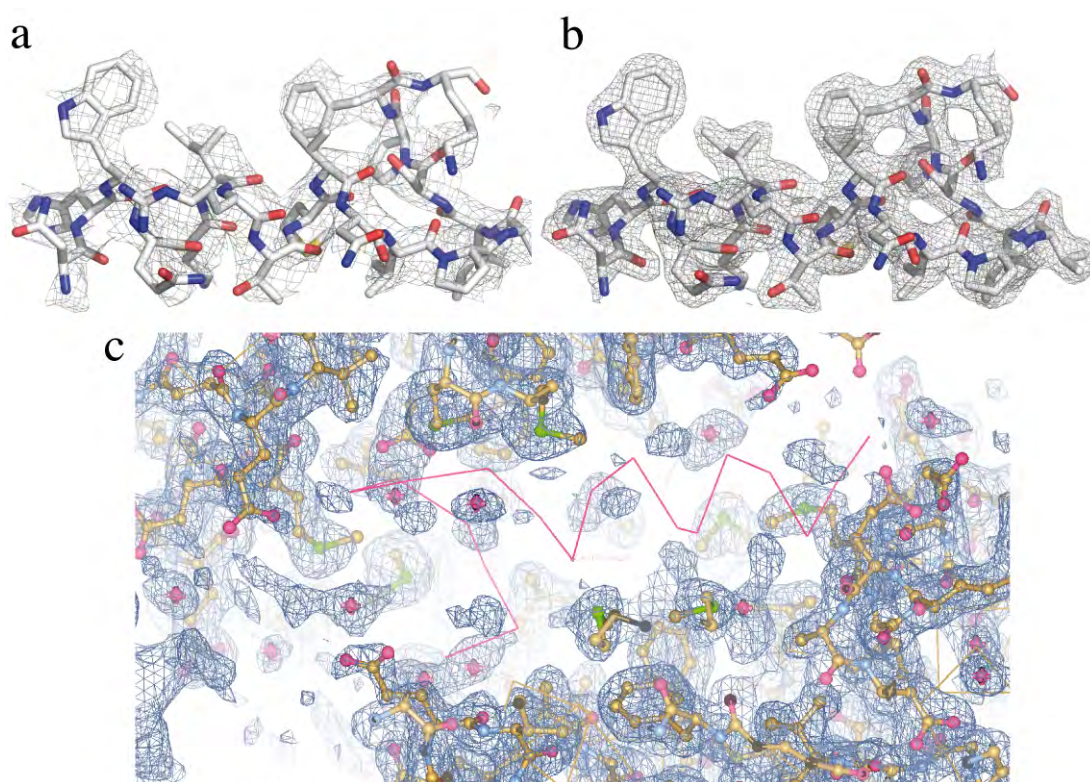
En cuanto a los cristales obtenidos empleando detergente, mostraron una resolución máxima de 1.84 Å (**Tabla 4.13**). Estos cristales presentaron un grupo espacial diferente a los cristales analizados anteriormente (P3<sub>2</sub>21;  $a = b = 40.27$  Å,  $c = 175.31$  Å) (**Figura 4.48e y f**). La estructura se determinó mediante Reemplazamiento Molecular (**Tabla 4.14**). La inspección del mapa de densidad reveló que estos cristales contenían solamente la molécula de CaM, no pudiéndose apreciar restos de densidad que indicasen la presencia del CBD de IP<sub>3</sub> 3-K (**Figura 4.49c**). Probablemente esto se debe a la presencia del detergente, el cual interfiere en las interacciones del CBD con CaM. La afinidad de muchos detergentes por CaM es igual, o incluso mayor, que la de muchos CBDs, como ya ha sido reportado (Kovacs *et al.*, 2010). Pese a no tratarse del complejo, se modeló y refinó la estructura de CaM obtenida (**Tabla 4.15**). La zona de mayor desorden está localizada en el bucle que une los dominios N y C-terminal de CaM.

Tabla 4.15. Estadísticas de refinamiento del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> y CaM.

	Hábito 1 + espermidina  (CaM/IP <sub>3</sub> 3-K <sub>(158-183)</sub> )	Hábito 2  (CaM)
<b>Refinamiento</b>		
Resolución (Å)	79.62-2.34	43.83-1.84
R <sub>work</sub> (%)	20.32	19.73
R <sub>free</sub> (%)	23.52	22.91
<b>Número de átomos (no H)</b>		
Totales	1428	1206
Proteína	1309	1108
Iones	14	4
Moléculas de agua	105	85
<b>Desviaciones cuadráticas medias (rmsd)</b>		
Enlaces (Å)	0.020	0.018
Ángulos (°)	1.90	1.74
<b>Factor térmico B(Å<sup>2</sup>)</b>		
Total	29.10	37.34
Proteína	29.46	36.95
<b>Diagrama de Ramachandran</b>		
Residuos en la zona más favorable (%)	94.5	95.9
<b>Residuos no modelados</b>		
	CaM: 1:3, 149 IP <sub>3</sub> 3-K <sub>(158-183)</sub> :158-164	CaM:1:3, 76:81, 148:149

$$R_{\text{work}}/R_{\text{free}} = \sum_{hkl} |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{calc}}| / \sum_{hkl} F_{hkl}^{\text{obs}}$$



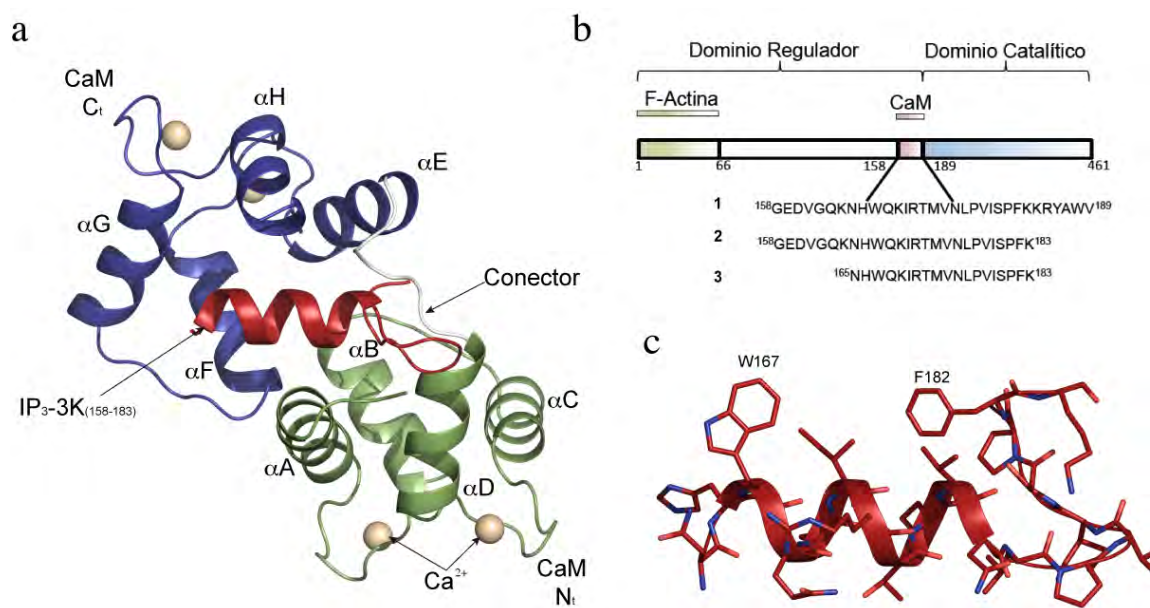


**Figura 4.49.** Mapas de densidad electrónica obtenidos de los diferentes cristales analizados. a) Mapa de densidad del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> a 1σ (3.8 Å) proveniente de los cristales obtenidos sin espermidina. b) Mapa de densidad del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> a 1σ (2.3 Å) proveniente de los cristales obtenidos con espermidina. c) Mapa de densidad obtenido a partir de los datos de cristales que contenían óxido de n-undecil-N,N-dimetilamina. Se ha superpuesto la estructura del complejo entre CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>. No se observa densidad electrónica en el sitio correspondiente a la hélice del CBD, lo que muestra que los cristales crecidos en presencia de óxido de n-undecil-N,N-dimetilamina están formados exclusivamente por moléculas de CaM y no del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.

#### 4.3.4 Estructura general del complejo

El complejo entre CaM e IP<sub>3</sub> 3-K<sub>(158-183)</sub> muestra una estructura compacta de forma elipsoidal y unas dimensiones aproximadas de 50x30x30 Å. Como se explicó en la Introducción, la CaM presenta varias conformaciones. En concreto la forma unida a CBDs es una forma cerrada de CaM, en la que los dos dominios de la CaM, N y C-terminal, (Yañez *et al.*, 2012) se encuentran abrazando el CBD. En la estructura del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> (**Figura 4.50a**) los dominios N y C-terminal de la CaM se encuentran en contacto a través del extremo C-terminal de la hélice αB (correspondiente al dominio N-terminal) y el extremo N-terminal de la hélice αF (correspondiente al dominio C-terminal). La estructura revela que, al igual que en otros complejos de este tipo, el péptido correspondiente a IP<sub>3</sub> 3-K<sub>(158-183)</sub> está unido a ambos dominios de la CaM (**Figura 4.50a**).



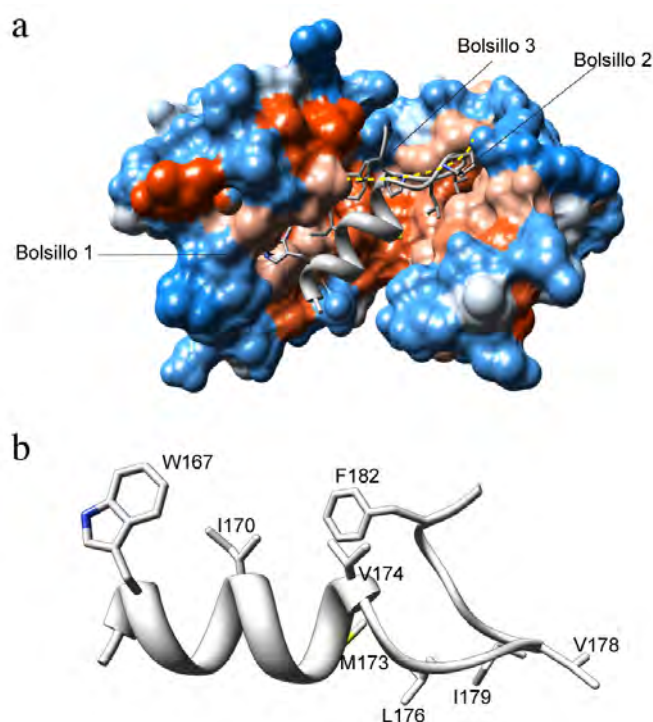


**Figura 4.50. Complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) Estructura general del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>. El dominio N<sub>i</sub> de CaM se muestra en color azul, en color blanco el bucle conector entre ambos dominios y en color verde el dominio C<sub>t</sub>. Las esferas de color crema representan átomos de Ca<sup>2+</sup>. IP<sub>3</sub> 3-K<sub>(158-183)</sub> se representa en color rojo. b) Esquema de la localización del dominio de unión a CaM dentro de IP<sub>3</sub> 3-K A. La secuencia número 1 representa la región mapeada como dominio de unión a CaM según Takazawa y Erneux, (1991). La secuencia número 2 representa el fragmento empleado en los experimentos de cristalización. La secuencia número 3 representa el fragmento visible en el mapa de densidad electrónica obtenido. c) Estructura del fragmento correspondiente a la secuencia 3.

La región de unión a CaM de IP<sub>3</sub> 3-K A empleada en el estudio estructural abarca los residuos 158-183 (**Figura 4.50b**). En la figura se puede apreciar densidad electrónica para el fragmento comprendido entre los residuos 165-183 (**Figura 4.50c**). Este fragmento se encuentra en el interior de un túnel de un diámetro de alrededor de 19 Å. Este túnel presenta una gran cantidad de residuos hidrofóbicos que permiten que el péptido se acomode perfectamente en su interior (**Figura 4.51a**). El área total de contacto entre ambas moléculas es de 1406.2 Å<sup>2</sup>.

El CBD de la IP<sub>3</sub> 3-K muestra dos zonas bien diferenciadas: una primera región situada en el extremo N-terminal, la cual se encuentra formando una corta hélice α de 10 residuos (residuos 165-175) con una longitud de 14.7 Å, y una segunda zona, en el extremo C-terminal la cual está formando un bucle en forma de horquilla (residuos 176-183), que se encuentra empaquetada contra la propia hélice α en su tramo final (residuos 182-183) (**Figura 4.50c**). El conjunto presenta una longitud total de 27.2 Å. La secuencia del CBD contiene un gran número de residuos con carácter hidrofóbico: Trp167, Ile170, Met173, Val174, Leu176, Val178, Ile179 y Phe182 (**Figura 4.51b**). Tanto la hélice α como la horquilla mantienen una red de interacciones con CaM (más de 130 contactos con distancias menores de 4 Å) siendo mayoritarios los contactos de

van der Waals (un 80% sobre el total de las interacciones).

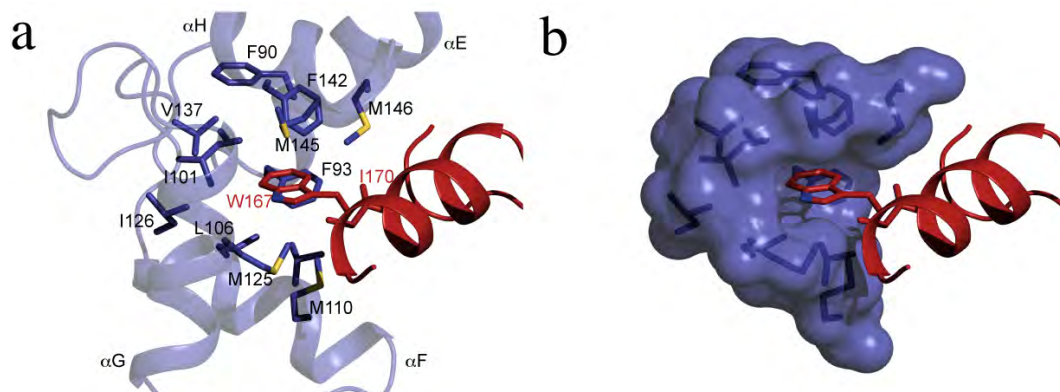


**Figura 4.51. Los residuos hidrofóbicos tienen un papel clave en la formación del complejo CaM/ IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) La CaM forma un túnel hidrofóbico en el cual se une el CBD. La CaM se representa en forma de superficie coloreada, correspondiendo los tonos rojos a residuos hidrofóbicos y los tonos azules a residuos no hidrofóbicos. Para mostrar con claridad el túnel se omitieron de la representación los aminoácidos correspondientes al conector entre los dominios de la CaM (residuos 77-82), representándose su posición mediante una línea discontinua amarilla. La IP<sub>3</sub> 3-K<sub>(158-183)</sub> se muestra en color blanco. La escala de hidrofobicidad para asignar los diferentes colores fue desarrollada por Eisenberg *et al.*, (1984). b) El fragmento IP<sub>3</sub> 3-K<sub>(158-183)</sub> (representado en color blanco) presenta también múltiples residuos hidrofóbicos (rotulados en la figura).

La mayor parte de las interacciones clave entre ambas moléculas se concentran alrededor de tres bolsillos bien diferenciados que pueden apreciarse en la CaM (**Figuras 4.52, 4.53 y 4.54**).

El primero de estos bolsillos está formado por residuos que provienen de las hélices  $\alpha$ E,  $\alpha$ F,  $\alpha$ G,  $\alpha$ H y el bucle que conecta las hélices  $\alpha$ G y  $\alpha$ H. Este bolsillo acoge los dos primeros residuos hidrofóbicos de la hélice de CBD, Trp167 e Ile170, ambos en la misma cara de la hélice  $\alpha$ . Dos residuos de metionina de CaM (Met125 y Met145) son los principales encargados de coordinar el Trp167, mientras que otro par de metioninas (Met110 y Met146) participan en la unión de Ile170. El resto del bolsillo se completa con otros residuos hidrofóbicos: Phe93, Ile101, Leu106, Val109, Ile126, Val137, Phe141 así como con Ala129 (**Figura 4.52a y b**). Los contactos de van der Waals entre

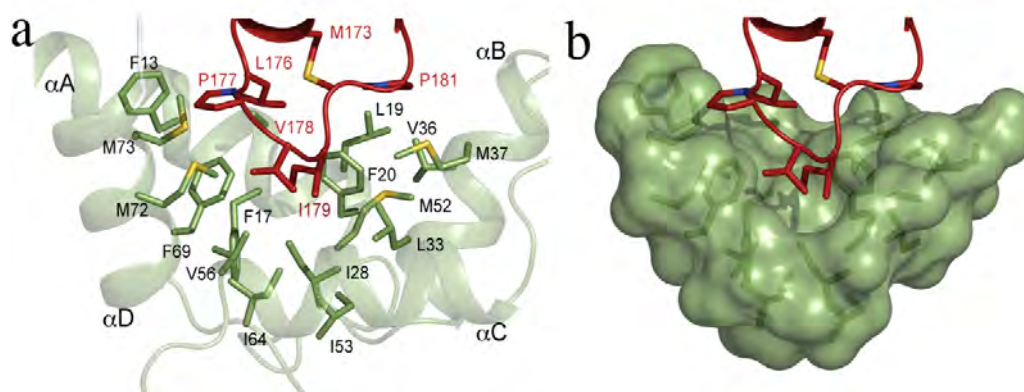
Trp167 y el par de metioninas Met110 y Met146 ayudan a anclar el extremo N-terminal del CBD al dominio C-terminal de la CaM. Este tipo de interacciones han sido observadas ya en otros complejos con CaM (Vetter y Leclerc, 2003) siendo habitual que este primer residuo hidrofóbico (normalmente un triptófano), el cual está totalmente conservado en los diferentes dominios de unión a CaM, se tome como referencia para definir los diferentes motivos de unión a CaM (Rhoads y Friedberg, 1997; Yap *et al.*, 2000).



**Figura 4.52. Bolsillo hidrofóbico 1 del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) Todos los residuos que conforman este bolsillo pertenecen al dominio C<sub>u</sub>, por lo que están representados en azul. El dominio de unión a CaM de IP<sub>3</sub> 3-K se representa en rojo. b) Cavidad formada por los diferentes residuos del dominio C<sub>t</sub> de CaM (superficie en color azul). La cavidad está ocupada por residuos de IP<sub>3</sub> 3-K (principalmente Trp167).

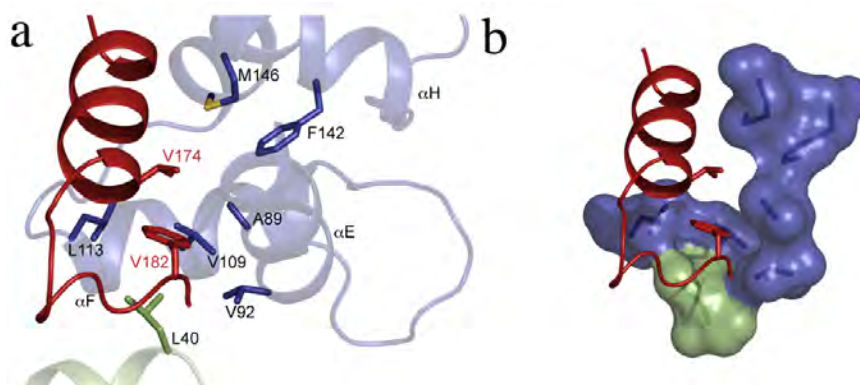
El segundo bolsillo hidrofóbico (**Figura 4.53a y b**), de gran tamaño y profundidad, está conformado por residuos que provienen de todas las hélices del dominio N-terminal de CaM y del bucle que conecta las  $\alpha$  hélices  $\alpha A$  y  $\alpha B$ . Este bolsillo acomoda la mayor parte del clúster de residuos hidrofóbicos localizados en la horquilla C-terminal del CBD: Leu176, Pro177, Val178, Ile179 y Pro181, así como la Met173. Las paredes del bolsillo están formadas por Phe13, Ala16, Phe17, Leu19 y Phe20 (hélice  $\alpha A$ ), Leu33, Val36 y Met37 (hélice  $\alpha B$ ), Met52 y Val56 (hélice  $\alpha C$ ), Met72 y Met73 (hélice  $\alpha D$ ). La región interior del bolsillo está formado por la Phe69 (hélice  $\alpha D$ ), Ile28 (bucle entre las hélices  $\alpha A$  y  $\alpha D$ ) e Ile64 (bucle entre las hélices  $\alpha B$  y  $\alpha C$ ). Los diferentes residuos hidrofóbicos de este clúster se encuentran dispuestos alrededor de las paredes de este bolsillo. Así pues, el primero de los residuos, Met173, está coordinado en la zona más abierta de este bolsillo por Ala16, Leu19, Phe20 y Met37. Leu176 es coordinada por un trío de fenilalaninas (Phe13, Phe20, Phe69) y Met73. Pro177 es la encargada de interaccionar con el par de Met situadas en la hélice  $\alpha D$  (Met72 y Met73). El siguiente residuo del CBD, Val178, está acomodado en una zona más amplia, formada por Met52, Val56 y Met72. Ile179 es el residuo más enterrado de todos los del clúster,

formando interacciones con un gran número de residuos del bolsillo: Phe20, Ile28, Leu33, Met37, Met51, Val56, Phe69 y Met71. Es muy probable que este residuo sea fundamental para el anclaje de la zona C-terminal del CBD, dado el gran número de interacciones con el dominio N-terminal de la CaM. Por último, Pro181 interacciona con Leu19, Phe20, Val36 y Met37.



**Figura 4.53. Bolsillo hidrofóbico 2 del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) Todos los residuos que conforman este bolsillo pertenecen al dominio N<sub>i</sub>, por lo que se representan en color verde. El dominio de unión a la CaM de IP<sub>3</sub> 3-K se representa en rojo. b) Cavidad formada por los diferentes residuos del dominio C<sub>i</sub> de la CaM (superficie en verde). Puede observarse como esta cavidad es ocupada por residuos de IP<sub>3</sub> 3-K.

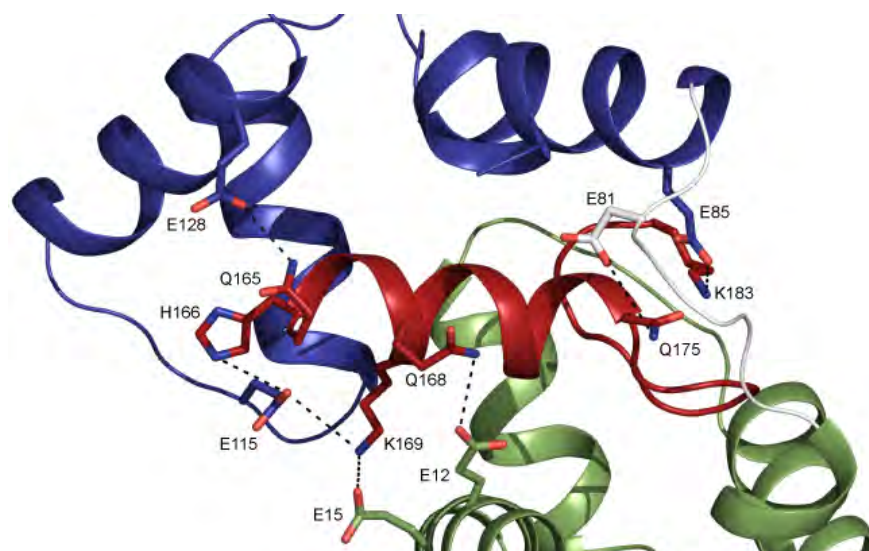
El tercero de los bolsillos (**Figura 4.54a y b**) acomoda la Val174 y Phe182 del CBD. Este bolsillo está formado por residuos procedentes de ambos dominios de la CaM: Leu40, Ala89, Val92, Val109 y Leu113. Este bolsillo, contiguo al bolsillo 1, se encuentra más abierto que los anteriores, siendo encargado de fijar el C-terminal de la horquilla, especialmente Phe182. El techo de la cavidad está formado por Phe142 y Met146, las cuales establecen interacciones hidrofóbicas con Val174.



**Figura 4.54. Bolsillo hidrofóbico 3 del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** a) Cavidad formada por residuos que provienen de ambos dominios de CaM (el código de colores es el mismo que en la **Figura 4.50**). b) Representación en superficie de la cavidad formada.



En cuanto a las interacciones polares (**Figura 4.55**), si bien no son muy numerosas, podrían ayudar a la estabilidad general del complejo y/o selección del CBD. Así pues, en el extremo N-terminal de la hélice del CBD, Gln165 se encuentra interaccionando con Glu128 (hélice  $\alpha$ G), His166 con Glu115 (bucle entre las hélices  $\alpha$ F/ $\alpha$ G) y Lys 169 con Glu15 (hélice  $\alpha$ A) y Glu115. Este conjunto de interacciones contribuyen a estabilizar el extremo N-terminal del CBD, junto con las interacciones hidrofóbicas que establece Trp167. Por otro lado, Gln168 interacciona con Glu12, Gln175 con Asp81 y Lys183 con Glu85. Esta última interacción en particular podría estar ayudando a orientar el extremo C-terminal del CBD hacia la salida del túnel hidrofóbico de CaM.



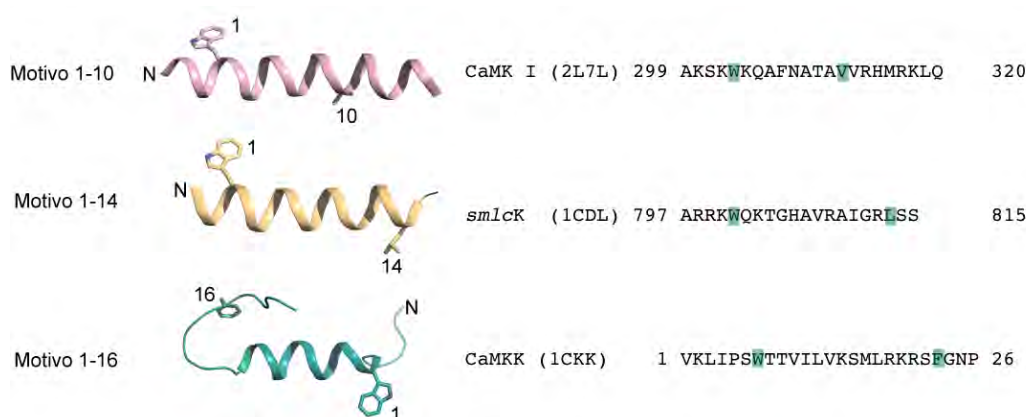
**Figura 4.55. Interacciones polares del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>.** Ambas moléculas (CaM y CBD) se representan mediante un modelo de cintas, detallándose los diferentes residuos que participan en las interacciones polares. El código de colores es el mismo que en la **Figura 4.50**. Las líneas discontinuas negras señalan las interacciones.

### 4.3.5 Comparación con otros motivos de unión a CaM (CBDs)

Como se ha visto, la CaM adquiere una conformación cerrada cuando está unida a CBDs, generándose varios bolsillos hidrofóbicos. Los CBDs suelen constar de 2 residuos principales, separados por una cierta distancia variable, siendo el residuo número uno el primero de los residuos hidrofóbicos del CBD comenzando en el N-terminal de este (comúnmente un Trp). Normalmente este primer residuo hidrofóbico del CBD suele estar interaccionando con el primer bolsillo hidrofóbico de la CaM (hélices en dirección N-terminal/C-terminal) salvo contadas excepciones, en las que al estar el CBD invertido (hélices en dirección C-terminal/N-terminal) este primer bolsillo está ocupado por el residuo situado en el extremo C-terminal del CBD. La clasificación de estos CBDs se basa en la distancia entre los residuos que ocupan el primer y segundo

bolsillo hidrofóbico de CaM (**Figura 4.56**).

La posición relativa de los diferentes residuos hidrofóbicos del dominio de unión a CaM son un aspecto clave para la caracterización de los diferentes CBD. Así pues, se ha comparado la estructura del complejo obtenido con representantes de tres motivos de unión caracterizados hasta la fecha (motivos 1-10, 1-14 y 1-16). Para ello, se ha utilizado como representante del motivo 1-10 el complejo entre  $\text{Ca}^{2+}$ /CaM y el CBD de la CaM kinasa (CaMK) I (código PDB 2L7L, (Gifford *et al.*, 2011)), como representante del motivo 1-14 el complejo entre la CaM y el CBD de la cadena ligera de miosina del músculo liso (sMLCK) (código PDB 1CDL, (Meador *et al.*, 1992)) y como representante del motivo 1-16 el complejo entre la CaM y el CBD de la CaMK Kinasa (código PDB 1CKK, (Osawa *et al.*, 1999)) (**Figura 4.56**).



**Figura 4.56. Motivos de unión a CaM.** A la izquierda se encuentra una representación del CBD de cada molécula, indicándose los residuos que definen el motivo, así como dónde está localizado el extremo N-terminal. A la derecha se muestra la secuencia de cada uno de ellos, recuadrándose en verde los residuos que definen el tipo de motivo, así como el nombre de la proteína a la que pertenecen y el código PDB correspondiente.

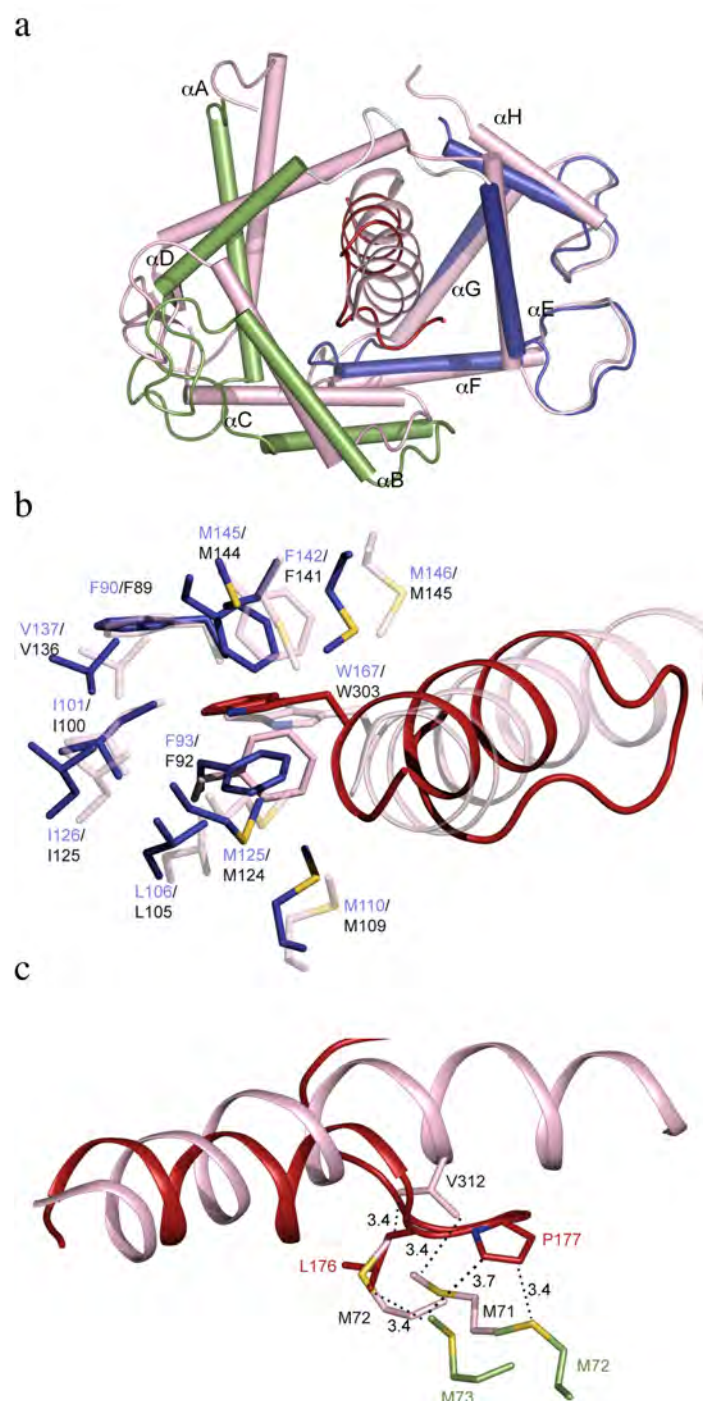
#### 4.3.5.1 Comparación con la CaMK I. Motivo tipo 1-10

La superposición de las estructuras de ambos complejos (CaM/IP<sub>3</sub> 3-K y CaM/CaMK I) revela que en ambos casos la CaM adopta una conformación cerrada, si bien ésta es diferente en ambas (**Figura 4.57a**). La superposición de los Cα tanto del dominio N-terminal (4-77) como del dominio C-terminal (77-82) presenta valores de desviación cuadrática media algo elevados, particularmente para la superposición del dominio N-terminal (1.50 Å<sup>2</sup> para el N-terminal, 0.79 Å<sup>2</sup> para el C-terminal y 1.95 Å<sup>2</sup> para la superposición de la estructura completa). Las diferencias en estos valores son debidas a la posición relativa de las diferentes hélices de CaM. En particular, puede observarse que el canal donde se encuentra unido el péptido es más ancho para el complejo entre CaM e IP<sub>3</sub> 3-K que para el formado entre CaM y CaM kinasa I (**Figura 4.57a**). Esto es debido a la diferente conformación de las hélices αA y αD, lo que

podría estar causado por la rotura de la hélice en el CBD de la IP<sub>3</sub> 3-K provocada por la presencia de Pro177. De esta manera, las hélices de la CaM de esta zona, al adoptar una conformación más abierta, pueden interactuar con el inicio de la horquilla. Es destacable también el hecho de que el bucle que conecta ambos dominios se encuentra mucho más desordenado en el caso del complejo con CaM/IP<sub>3</sub> 3-K. En el caso del dominio C-terminal pueden apreciarse también cambios en las hélices, si bien son mucho menos significativos, como cabe esperar tras observar las interacciones conservadas del primer residuo de anclaje en ambos complejos.

En cuanto a la comparación entre los CBDs, puede observarse que la hélice formada por el CBD es mucho más larga en el caso de la CaM kinasa I (26.4 Å frente a 14.7 Å en el caso de IP<sub>3</sub> 3-K) (**Figura 4.57**). En ambos casos el primer residuo de anclaje es un Trp (Trp303), el cual interactúa con el dominio N-terminal de la CaM. La posición de este Trp303 es equivalente en ambos complejos, presentando además el mismo rotámero. El triptófano está ocluido en un bolsillo de unión formado por los mismos residuos en ambos casos, si bien existen diferencias en las conformaciones de algunos de ellos. Estas diferencias, visibles principalmente en las metioninas, reflejan los ajustes locales que se producen en cada caso, dado que la posición de la hélice no es exactamente igual en ambos complejos (**Figura 4.57b**). En el caso del segundo bolsillo, puede observarse que el residuo que lo ocupa en la CaM kinasa I es un residuo de valina (Val312). Este está separado de este primer Trp por 9 residuos, siendo por tanto un motivo de unión de tipo 1-10. En caso del dominio de unión de la IP<sub>3</sub> 3-K A, el segundo residuo equivalente corresponde a la Leu176 (**Figura 4.57c**). En la CaM Kinasa I, Val312 interactúa con el par de metioninas formado por Met71 y Met72. En el caso del CBD de IP<sub>3</sub> 3-K, Leu176 solo interactúa con una de estas metioninas (Met73, equivalente a Met72 en la CaM Kinasa I). La segunda de estas metioninas (Met74, equivalente a Met72 en la CaM Kinasa I) se encuentra muy alejada de la Leu172. Sin embargo, en IP<sub>3</sub> 3-K sí existe un residuo que realiza interacciones similares con ambas metioninas: Pro177, si bien por su posición (número 11 respecto al Trp165) y disposición espacial dentro del complejo no parece cumplir ninguna función de “anclaje” (**Figura 4.57c**). Finalmente, el tercer bolsillo hidrofóbico descrito para el complejo entre CaM e IP<sub>3</sub> 3-K no está presente en el complejo entre CaM y CBD de la CaM Kinasa I. Las interacciones polares descritas en el complejo CaM/IP<sub>3</sub> 3-K tampoco se encuentran conservadas en el complejo CaM/CaM Kinasa.





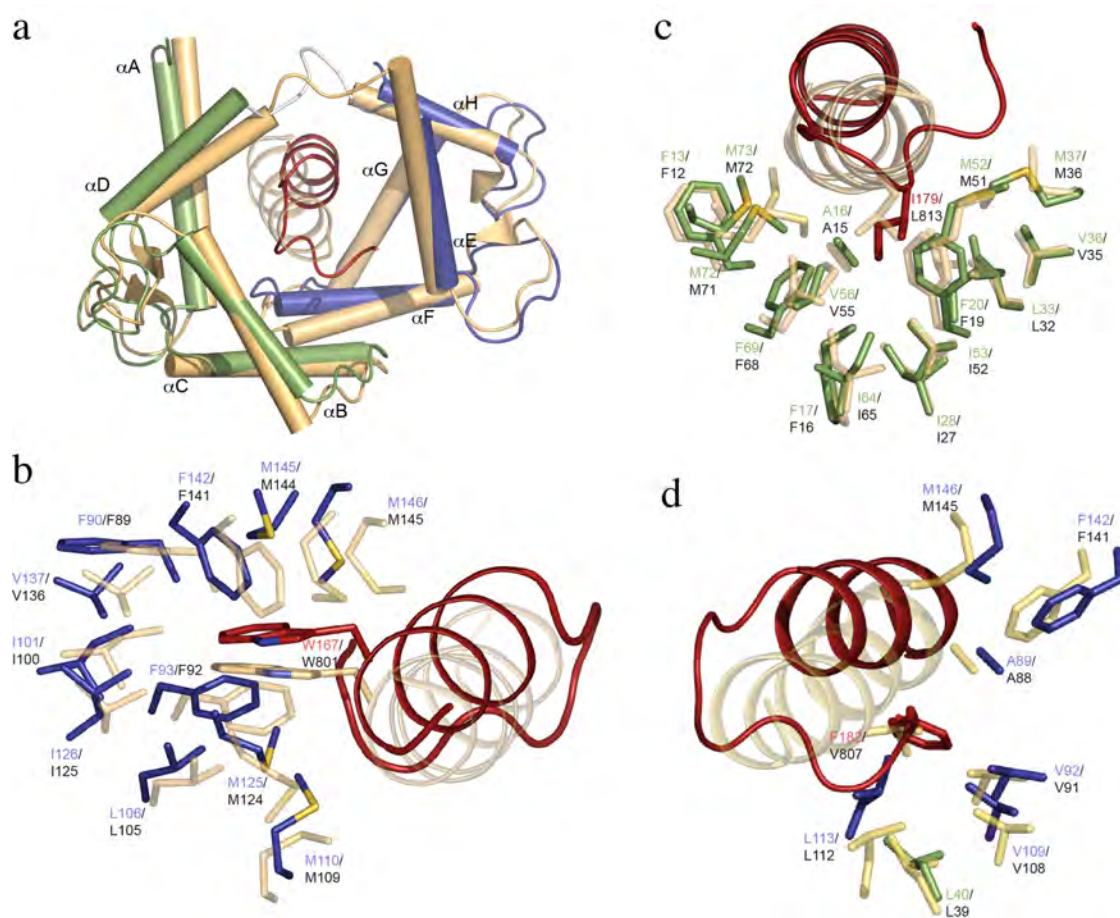
**Figura 4.57. Comparación del CBD de IP<sub>3</sub> 3-K con un motivo de tipo 1-10.** El código de colores para la estructura de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> es el mismo que en la Figura 3.57 y siguientes. La estructura de CaM/CBD de CaMK I se representa en color rosa. La numeración de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> sigue el código de colores explicado, mientras que la numeración en negro corresponde al complejo CaM/CaMK I. a) Vista general de ambas estructuras superpuestas. Puede observarse que el complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> se encuentra en una conformación más abierta. b) Comparación del bolsillo 1. Puede apreciarse la gran similitud entre ambos bolsillos. c) Detalle de las interacciones del residuo en posición 10 de ambos complejos (bolsillo 2). En el caso de IP<sub>3</sub> 3-K es el residuo en posición 11 (Pro177) y no el residuo en posición 10 (Leu 176) el que realiza interacciones equivalentes a las que realiza el residuo 10 (Val312) en el complejo CaM/CaM kinasa I.

#### 4.3.5.2 Comparación con la sMLCK. Motivo tipo 1-14

La conformación de la CaM es también ligeramente diferente en los complejos con CBDs de la IP<sub>3</sub> 3-K o de la sMLCK, como puede apreciarse al superponer las estructuras correspondientes (rmsd = 0.515 Å<sup>2</sup> para el dominio N-terminal y rmsd = 0.710 Å<sup>2</sup> para el dominio C-terminal y rmsd = 1.517 Å<sup>2</sup> para ambos dominios). En este caso, la CaM en el complejo con IP<sub>3</sub> 3-K se encuentra también en una conformación más abierta. Las mayores diferencias se localizan en el dominio N-terminal, adoptando las hélices  $\alpha A$  y  $\alpha D$  una conformación que amplía el canal que acoge el CBD (**Figura 4.58a**). Por otro lado, el dominio C-terminal presenta una conformación muy similar en ambos complejos. Las mayores diferencias se sitúan en el bucle de conexión entre ambos dominios de la CaM. Buena parte de este bucle se encuentra desordenado en el complejo entre CaM e IP<sub>3</sub> 3-K, mientras que en el complejo entre CaM/sMLCK se encuentra ordenado formando parte de la hélice  $\alpha E$ .

La hélice del CBD es más larga en el caso del complejo CaM/sMLCK (27.4 Å frente a 14.7 Å en la IP<sub>3</sub> 3-K). Al igual que en IP<sub>3</sub> 3-K y en el motivo tipo 1-10, el primer residuo de anclaje de la sMLCK corresponde a un triptófano (Trp801), el cual presenta el mismo rotámero en ambos complejos. De nuevo, este residuo está alojado en una cavidad hidrofóbica de similares características (bolsillo 1). Igualmente, los residuos del bolsillo en la CaM que más varían su conformación son los residuos de metionina, los cuales son los mas implicados en la interacción con el triptófano (**Figura 4.58b**). En cuanto al segundo de los residuos clave, en el caso de la sMLCK, éste corresponde a una Leu (Leu813). En el caso de la IP<sub>3</sub> 3-K, el residuo equivalente en secuencia es la Ser180, aunque este residuo no parece presentar interacciones relevantes para el anclaje a CaM. Sin embargo, el residuo precedente, Ile179, que corresponde con la posición 13, ocupa una posición estructural similar a la de la Leu813 (**Figura 4.58c**). El complejo CaM/sMLCK también presenta un residuo alojado en el tercer bolsillo hidrofóbico, si bien este residuo difiere en posición en la secuencia. Así pues, en el caso del complejo CaM/sMLCK este bolsillo está ocupado por Val807 (residuo en posición 8, dado que este dominio de unión a CaM corresponde a la subclase 1-8-14), mientras que en el caso del complejo entre CaM e IP<sub>3</sub> 3-K se encuentra ocupada por Phe183 (la cual se encuentra en la posición 16). Por lo tanto, puede observarse que se conserva el anclaje a los tres bolsillos descritos en la CaM, si bien a excepción del triptófano, alojado en el primer bolsillo, los residuos reconocidos por la CaM ocupan posiciones muy diferentes en ambos CBDs.

El patrón de interacciones polares es diferente, aunque en ambos casos el mayor número de interacciones de este tipo se concentran en el N-terminal de ambos CBDs.



**Figura 4.58. Comparación del motivo de unión a CaM de IP<sub>3</sub> 3-K y un motivo de tipo 1-14.** El código de colores para la estructura de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> es el mismo que en la Figura 3.57 y siguientes. La estructura de CaM/sMLCK se representa en color amarillo. La numeración de cada estructura sigue el código de colores explicado en la **Figura 4.57**, representándose en negro la numeración de CaM/sMLCK. a) Vista general de ambas estructuras superpuestas. Al igual que antes, el complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> se encuentra en una conformación más abierta. b) Comparación del bolsillo 1 de los dos complejos. Puede apreciarse la gran similitud entre ambos bolsillos. c) Detalle del bolsillo 2. El bolsillo es muy similar y en ambos casos está ocupado por un residuo hidrofóbico, con la diferencia de que en el caso de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> el residuo (Ile179) se encuentra en posición 13, en lugar de la 14. d) Detalle del bolsillo 3. El bolsillo se encuentra bastante conservado en ambos complejos, pese a que la posición del residuo que lo ocupa difiere (residuo en posición 8 en el caso de CaM/sMLCK y residuo en posición 16 en el caso de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>).

#### 4.3.5.3 Comparación con la CaMKK. Motivo tipo 1-16

La superposición de ambos complejos revela diferencias sustanciales entre ambos (rmsd = 0.651 Å<sup>2</sup> para el dominio N-terminal, rmsd = 1.014 Å<sup>2</sup> para el dominio C-terminal y rmsd = 1.830 Å<sup>2</sup> para la proteína completa). Las diferencias entre dominios son probablemente debidas a la diferente orientación de la hélice, la cual determina la posición de la horquilla C-terminal. Puede observarse que el dominio N-terminal se encuentra en una conformación más abierta en el complejo entre CaM/IP<sub>3</sub> 3-K (**Figura**

##### 4.59a).

El complejo entre CaM/CaMK Kinasa es el único cristalizado hasta la fecha que presenta una estructura de dominio de unión a CaM similar a la observada para el dominio de unión a CaM de IP<sub>3</sub> 3-K, es decir, una hélice  $\alpha$  y una horquilla. Si bien la hélice de IP<sub>3</sub> 3-K es de menor tamaño (14.7 Å frente 17.5 Å de la hélice de CaMKK, estando la hélice formada por 11 y 12 residuos respectivamente) la extensión global del motivo es mayor en IP<sub>3</sub> 3-K (27.2 Å frente a 24.3 Å en CaMKK).

A pesar de esta similitud, la unión es bien distinta, ya que la orientación de ambas hélices es opuesta: en el caso del complejo CaM/CaMKK el extremo N-terminal de la hélice CBD interacciona con el dominio N-terminal de CaM, mientras que en el caso del complejo CaM/IP<sub>3</sub> 3-K el extremo N-terminal de la hélice CBD interacciona con el dominio C-terminal de CaM (**Figura 4.60**).

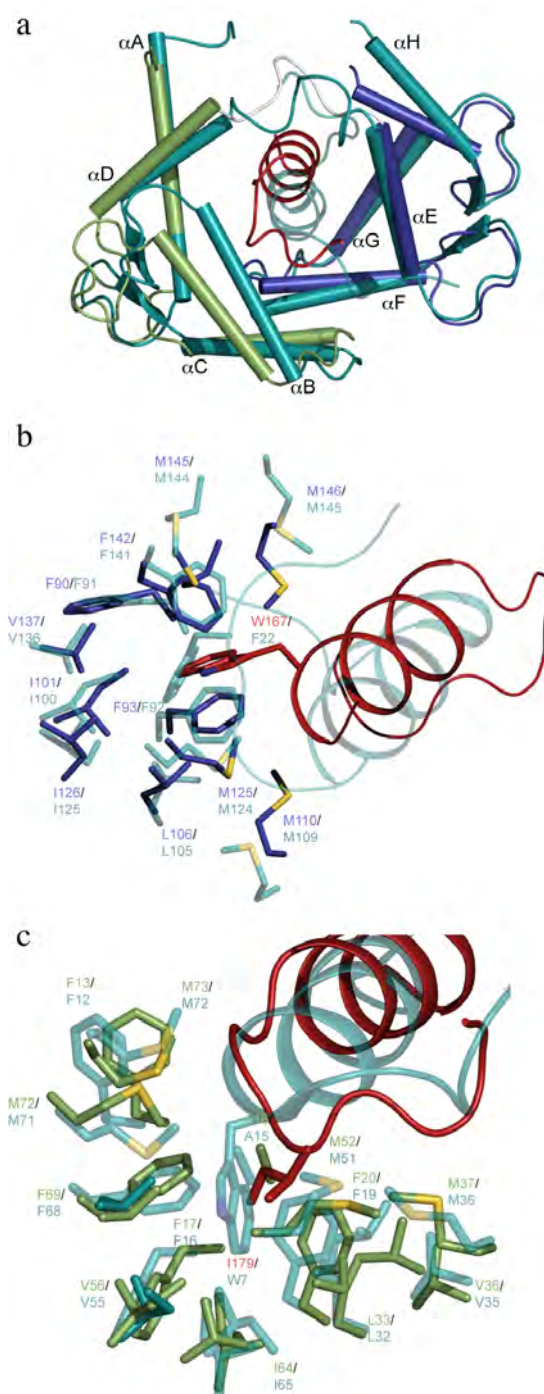
Así pues, el primer bolsillo hidrofóbico, localizado en el dominio C-terminal de la CaM, está ocupado en ambos casos por un residuo hidrofóbico de gran tamaño. En el caso de la IP<sub>3</sub> 3-K es el Trp167 (residuo 1). Sin embargo, en el caso de la CaMKK, este residuo es la Phe22 (residuo 16) (**Figura 4.59b**). Este primer bolsillo presenta una conformación ligeramente diferente en ambos casos, siendo una cavidad más abierta en el caso de la CaMKK (**Figura 4.59b**). Esto es debido a que la hélice  $\alpha$ H en el complejo CaM/CaMKK se encuentra desplazada debido a la presencia de la horquilla C-terminal de la CaMKK.

En el caso del segundo bolsillo de CaM, este se encuentra ocupado por el Trp7 en la CaMKK (residuo 1) y la Ile179 (residuo 13) en el caso de la IP<sub>3</sub> 3-K. Ambos residuos se encuentran ocupando la misma posición, ocluidos en el bolsillo hidrofóbico formado por residuos procedentes de las diferentes hélices del dominio C-terminal de CaM (**Figura 4.59c**).

En cuanto al tercer bolsillo hidrofóbico, ambos CBDs presentan un residuo de gran tamaño en el extremo C-terminal de la horquilla (Phe182 en la IP<sub>3</sub> 3-K y Phe22 en la CaMKK), si bien dada la diferente orientación de la hélice estos residuos realizan interacciones con diferentes zonas de la CaM. En ambos CBDs estos residuos de fenilalanina están precedidos de residuos de prolina, lo que muy probablemente ayuda a fijar la posición de estos residuos. En este caso tampoco se encuentran interacciones polares equivalentes entre ambos complejos.

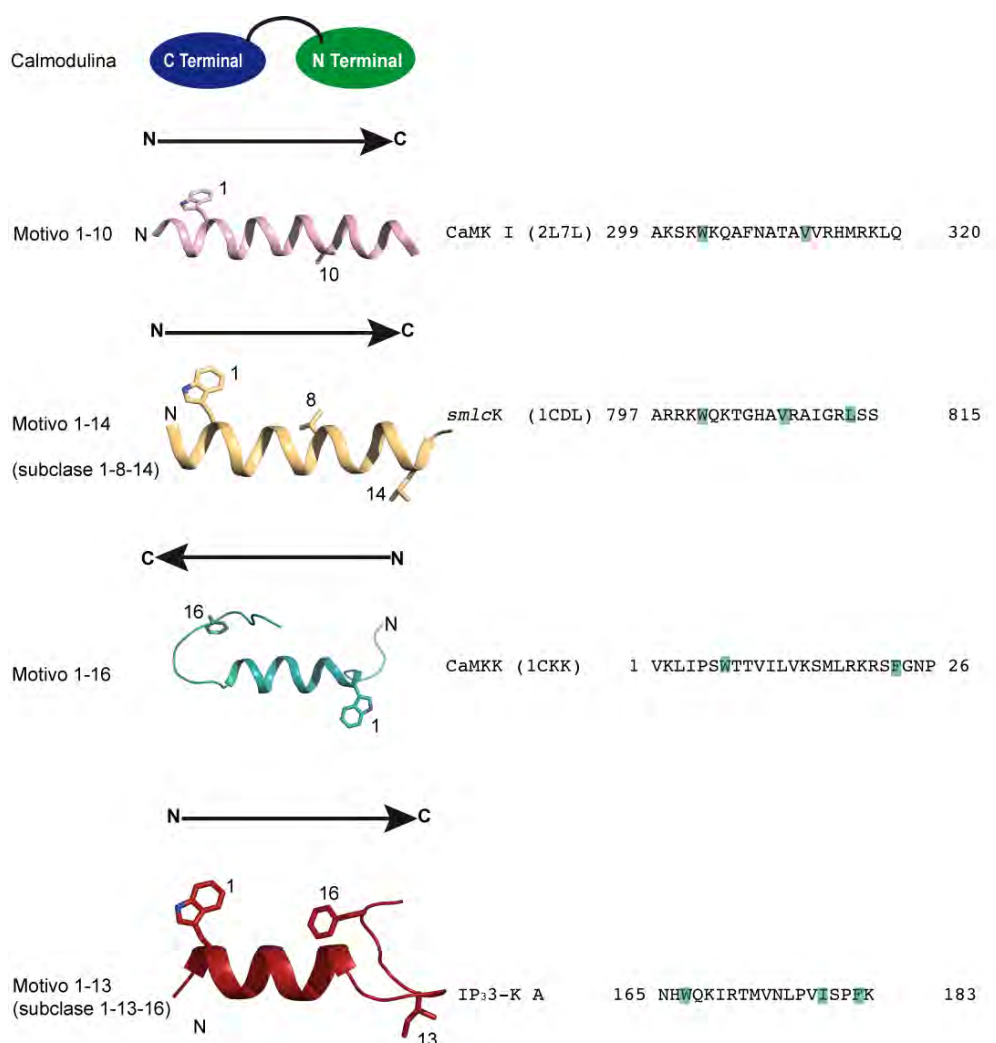
En cualquier caso, es destacable el gran parecido global entre la conformación y organización general de ambos CBDs.





**Figura 4.59. Comparación del motivo de unión a CaM de IP<sub>3</sub> 3-K y un motivo de tipo 1-16.** El código de colores para la estructura de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> es el mismo que en la Figura 3.57 y siguientes. La estructura del complejo CaM/CaMKK se representa en color cian. Las etiquetas siguen el mismo código de colores. a) Superposición de ambos complejos. b) Comparación del bolsillo 1 de ambos complejos. Este bolsillo es similar en ambos, pero está ocupado por el residuo 1 en el caso de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> y por el residuo 16 en el caso de CaM/CaMKK, debido a que las hélices se orientan en direcciones opuestas en cada uno de los complejos. c) Comparación del bolsillo 2. En este caso los residuos que ocupan el bolsillo son el residuo 13 (Ile179 en el complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>) y el 1 (Trp7 en el complejo CaM/CaMKK).

Teniendo en cuenta los análisis expuestos, el motivo de unión a la CaM de la IP<sub>3</sub> 3-K A no pertenece a ninguna de las tres clases previamente caracterizadas. Considerando la interacción de IP<sub>3</sub> 3-K con los bolsillos 1 y 2 de la CaM, puede decirse que el CBD estudiado define una nueva clase de motivo, el motivo 1-13. Si además consideramos la fuerte unión del residuo en posición 16 (Phe182) en el bolsillo 3 podría definirse una subclase 1-13-16. No obstante, habría que comprobar la interacción de estos sitios mediante otras técnicas, como por ejemplo mutagénesis dirigida.



**Figura 4.60.** El motivo de unión a CaM de la IP<sub>3</sub> 3-K A se encuadra dentro de una nueva clase de motivos 1-13-16. Además de esto se indica la dirección de los diferentes motivos de unión a CaM. La figura muestra en qué dirección está orientada la hélice del CBD correspondiente respecto a los dominios N y C-terminal de CaM. El código de colores es el mismo que en las Figuras 4.49 y 4.55.

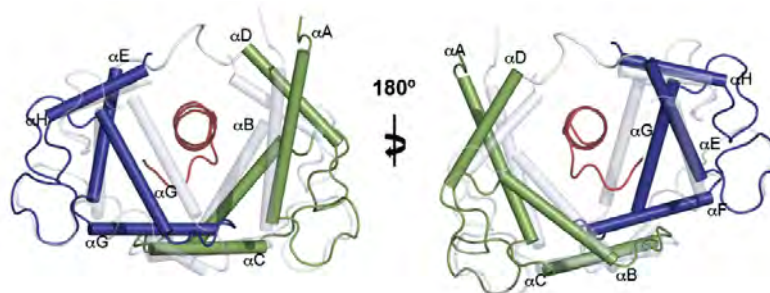
### 4.3.6 Comparación entre las dos formas de CaM observadas en este trabajo

Como se ha explicado anteriormente en la Sección 4.3.2, se obtuvieron también cristales de otro hábito en condiciones similares pero en presencia de un detergente (óxido de n-N,N-dimetilamina). Pese a que estos cristales eran de excelente calidad, no contenían el complejo entre CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub>, sino solamente una molécula de CaM. Es probable que el uso del detergente provocara la disociación del complejo, compitiendo con las interacciones hidrofóbicas que se forman.

Posiblemente los detergentes empleados son capaces de unirse a la cavidad hidrofóbica formada por los dos dominios de la CaM. Hasta la fecha solo hay una estructura de CaM en complejo con detergentes depositada en el PDB (código PDB 3if7, (Kovacs *et al.*, 2010). Al comparar esta estructura las aquí presentadas, se observa que la conformación es muy similar (rmsd = 0.459 Å<sup>2</sup> considerando ambos dominios) pero, mientras que en los cristales de la CaM de la estructura depositada se han encontrado cuatro moléculas de lípido, en los cristales obtenidos en este trabajo no se observó la presencia de moléculas de detergente. Es probable que esto se deba a la diferente naturaleza del detergente usado (esfingofosforilcolina en el caso del PDB 3if7; óxido de n-N,N-dimetilamina en el caso de los cristales obtenidos en este trabajo).

La CaM cristalizada en presencia de detergente presenta también una conformación cerrada, pero muestra notables diferencias respecto a la conformación del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> (rmsd = 3.589 Å<sup>2</sup> para un total de 133 residuos), aunque la conformación de los diferentes dominios no es tan diferente entre sí (rmsd = 0.500 Å<sup>2</sup> para el dominio N-terminal y rmsd = 0.728 Å<sup>2</sup> para el dominio C-terminal). La estructura reportada representa una forma más cerrada de CaM, con un diámetro del túnel hidrofóbico significativamente menor que en el caso del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> (**Figura 4.61**). Además este túnel se encuentra abierto en su zona superior, debido a que en esta conformación el conector entre los dominios N y C de CaM se encuentra desordenado (probablemente al desaparecer las interacciones con IP<sub>3</sub> 3-K<sub>(158-183)</sub> el bucle pasa a una conformación más flexible) (**Figura 4.61**). Pueden apreciarse cambios significativos en la mayor parte de las hélices: αA, αD, αE, αF, αG y αH, apreciándose un desplazamiento de cada una de estas hélices de hasta 4 Å, lo que genera un túnel más abierto donde puede acomodarse el dominio de unión a CaM. Existe también una reorganización notable en los diferentes bucles que conectan las hélices, especialmente en aquellos que conectan las hélices αC/αD, αE/αF, αF/αG y αG/αH.



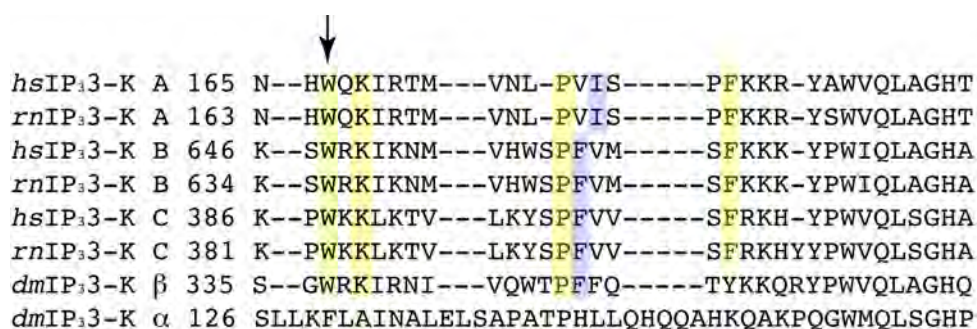


**Figura 4.61. Comparación entre las dos formas de CaM observadas en este trabajo.** Las estructuras del complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> y de la CaM obtenida mediante el uso de óxido de n-undecil-N,N-dimetilamina han sido superpuestas. El código de colores para la estructura de CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> es el mismo que en la Figura 3.57 y siguientes, mientras que la estructura de la CaM obtenida con óxido de n-undecil-N,N-dimetilamina se representa en color blanco. Puede observarse que el complejo CaM/IP<sub>3</sub> 3-K<sub>(158-183)</sub> presenta una conformación significativamente más abierta.

En conjunto, las distintas estructuras observadas son coherentes con la bien conocida flexibilidad intrínseca de CaM, gracias a la cual es capaz de adaptar su superficie a cualquier tipo de ligandos de naturaleza hidrofóbica (Yamniuk y Vogel, 2004).

#### 4.3.7 Comparación con otras isoformas de la IP<sub>3</sub> 3-K. Correlación con estudios bioquímicos

Un alineamiento de las posibles regiones de unión a la CaM de la IP<sub>3</sub> 3-K de diferentes especies reveló la conservación del motivo 1-13-16 caracterizado en este trabajo (algunas de las secuencias mas representativas se muestran en la **Figura 4.62**). Como puede observarse, el Trp clave para el anclaje del dominio de unión a CaM está conservado en todas las secuencias, salvo en la enzima de *C. elegans*, la cual no es activada mediante CaM, lo que concuerda con la ausencia de este residuo (Kucharski y Maleszka, 2002). El resto de residuos conservados son Lys169, Pro177 y Phe182. La Lys169 podría ser, por su posición, importante para la estabilización del Trp167 (**Figura 4.55** y **Figura 4.62**). Por otra parte, la conservación de una prolina en la posición de Pro177 (**Figura 4.53** y **Figura 4.62**), revela que es muy probable que la configuración de los dominios de unión a CaM de las diferentes IP<sub>3</sub> 3-Ks sea la misma, presentando todas ellas una hélice corta y una horquilla en el extremo C-terminal. Por su parte, la absoluta conservación de Phe182 (posición 16) indica que la interacción de este residuo con la CaM podría resultar importante para orientar el extremo C-terminal del dominio de unión hacia la salida del túnel (**Figura 4.54** y **Figura 4.62**). En cuanto al resto de residuos, si bien no están estrictamente conservados, sí está conservado su carácter (básico o hidrofóbico), existiendo residuos del mismo tipo en posiciones equivalentes. Así ocurre por ejemplo con el residuo en posición 13.



**Figura 4.62.** Alineamiento entre las diferentes isoformas de la IP<sub>3</sub> 3-K de diferentes especies. *hs*: *H. sapiens*; *rn*: *R. norvegicus*; *dm*: *D. melanogaster*. El sombreado amarillo indica los residuos conservados. La flecha indica la posición del triptófano que ocupa la posición 1 del motivo de unión a CaM, mientras que el sombreado violeta indica la posición del residuo número 13 en cada una de las secuencias. El residuo número 16 sería la última fenilalanina conservada en la secuencia.

Se han realizado diferentes estudios bioquímicos del dominio de unión a CaM de la IP<sub>3</sub> 3-K. Los primeros, realizados con la isoforma A, fueron encaminados a acotar el dominio propiamente dicho, estando comprendido entre los residuos 156-189 (Takazawa y Erneux, 1991). El trabajo aquí presentado confirma que el dominio de unión a CaM se encuentra dentro de esa región, si bien los residuos de los segmentos 156-164 y 184-189 no parecen participar en la unión. Se han realizado también diferentes estudios de mutagénesis dirigida del CBD (Erneux *et al.*, 1993; Chamberlain *et al.*, 2005; Lloyd-Burton *et al.*, 2007). En ellos, el Trp165 de la isoforma A de rata fue mutado por glicina (W165G) (Erneux *et al.*, 1993) o arginina (W165R) (Erneux *et al.*, 1993; Lloyd-Burton *et al.*, 2007). Se han mutado también los triptófanos de las isoformas B (W636R) de rata (Lloyd-Burton *et al.*, 2007) y ratón (W644A) (Chamberlain *et al.*, 2005) y de la isoforma C (W383R) de rata (Lloyd-Burton *et al.*, 2007), así como el residuo equivalente en la isoforma β de *D. melanogaster* (W337R) (Lloyd-Burton *et al.*, 2007). Todos estos estudios confirman el papel crucial para la unión del CBD del triptófano en la posición 1 del motivo de unión, pues en todos los casos se observó que la mutación impedía la unión de la IP<sub>3</sub> 3-K. En aquellos casos en los que se midió la activación de la enzima por la CaM (Chamberlain *et al.*, 2005; Lloyd-Burton *et al.*, 2007) no se apreciaron diferencias de actividad respecto a la actividad basal de la enzima nativa, lo que sugiere que dichos mutantes han perdido su capacidad de activación por la CaM.

Los alineamientos de secuencia realizados hasta la fecha indicaban que probablemente el motivo de unión a CaM de la IP<sub>3</sub> 3-K pertenecía a la clase 1-10 (Chamberlain *et al.*, 2005). Con la intención de confirmar esta idea, se realizó la mutación del residuo en esta posición en la isoforma B de ratón por alanina (W653A). La mutación de este residuo mostró que su ausencia no variaba la activación de la IP<sub>3</sub> 3-K por CaM, lo que indica que este no es un residuo crítico para la unión de la CaM (Chamberlain *et al.*, 2005). Como se ha propuesto en este trabajo, muy probablemente

son los residuos en posición 13 y 16 los responsables del anclaje del C-terminal del CBD a CaM.

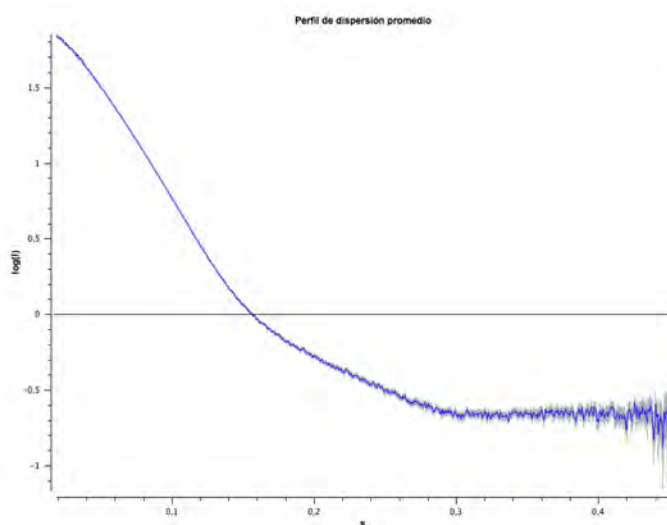
### 4.3.8 Un modelo del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub> mediante la combinación de datos a alta y baja resolución

Pese a ser una enzima clave para el flujo de Ca<sup>2+</sup> en la célula, se conocen pocos detalles acerca de cómo IP<sub>3</sub> 3-K es regulada por CaM. Con el objetivo de profundizar en esto se realizó un modelo del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub> mediante una combinación de datos de dispersión a bajo ángulo (SAXS) y cristalografía.

El perfil de dispersión promedio obtenido a partir de muestras de CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub> puede observarse en la **Figura 4.63**. Los datos sobre la partícula en solución obtenidos empleando SAXS pueden verse en la **Tabla 4.16**

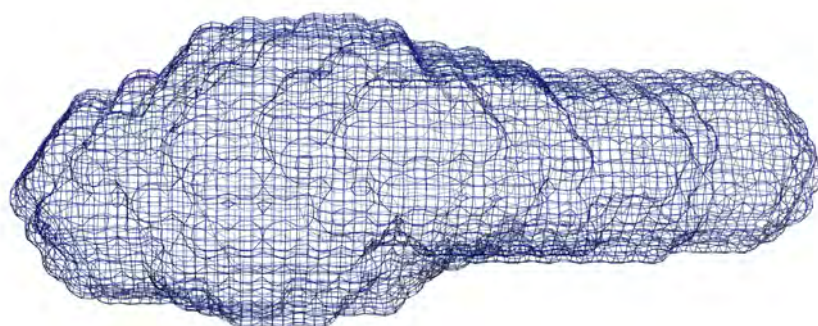
**Tabla 4.16. Parámetros estructurales obtenidos mediante SAXS.**

Parámetros estructurales	
<b>Parámetros estructurales</b>	
$I(0)$ (calculado según $P(r)$ )	87.68
$R_g$ (calculado según $P(r)$ )	3.97
$I(0)$ (calculado según <i>Guinier plot</i> )	86.55
$R_g$ (calculado según <i>Guinier plot</i> )	3.95
<b>Determinación de masa molecular</b>	
Masa molecular estimada (según $I(0)$ ) (Da)	87335
Masa molecular estimada (según secuencia) (Da)	55176



**Figura 4.63.** Perfil de dispersión promedio obtenido a partir de muestras del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub>. Los datos fueron cortados a  $s=0.35$ .

De estos datos se obtuvieron diferentes modelos de la partícula a baja resolución, los cuales fueron promediados. El promediado de estos modelos se muestra en la **Figura 4.64**. La envoltura obtenida presenta unas dimensiones aproximadas de 125 Å x 60 Å x 45 Å y se puede describir como un cuerpo central, de mayor tamaño y, en un extremo, una protuberancia más estrecha. Es probable que la parte más estrecha corresponda a la zona desestructurada del complejo (residuos 128-164) en el extremo N-terminal de la IP<sub>3</sub> 3-K.

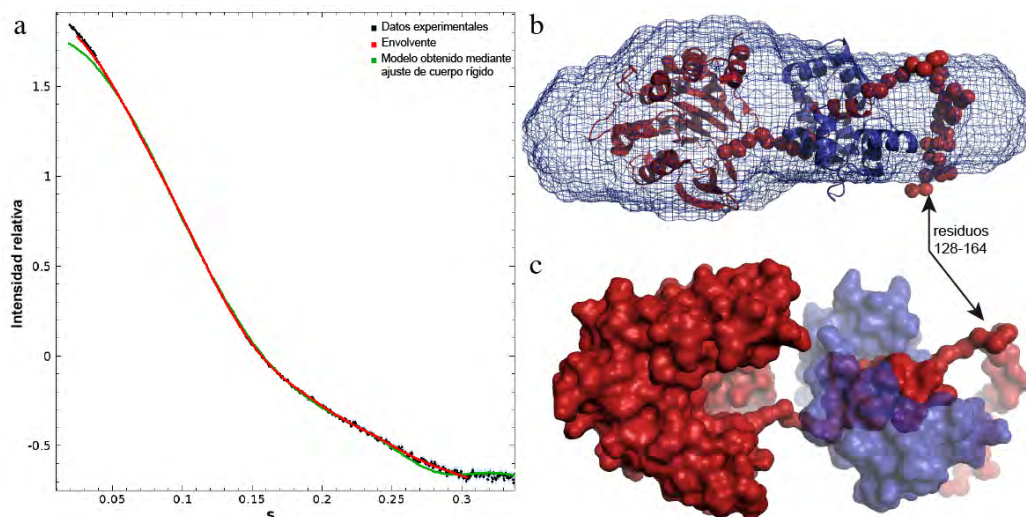


**Figura 4.64.** Envoltura obtenida a partir de los datos de dispersión a bajo ángulo (SAXS) del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub>. Este modelo muestra la forma aproximada de dicho complejo.

El modelo obtenido a baja resolución se combinó con las estructuras cristalográficas disponibles (complejo entre CaM/IP<sub>3</sub> 3-K<sub>(156-183)</sub> A y dominio catalítico de IP<sub>3</sub> 3-K A, el cual comprende los residuos 189-461) (González *et al.*, 2004). El modelo se elaboró mediante ajuste de cuerpo rígido de estas estructuras en el modelo de dispersión a bajo ángulo obtenido. Con el objetivo de intentar obtener un modelo más completo, se realizó un modelado de los conectores entre los dominios kinasa y CBD (fragmento 184-189) y del extremo N-terminal de la construcción de IP<sub>3</sub> 3-K utilizada (residuos 128-164). El ajuste entre los datos experimentales, el modelo a baja resolución y el modelo a alta resolución puede observarse en las **Figura 4.65a y b**. Para ello, se calculó la dispersión teórica en solución (*scattering* teórico) de los datos cristalográficos, tomándose el valor  $\chi^2$  como un indicador relativo de la bondad del ajuste entre el modelo y los datos. Valores bajos (idealmente <2, aunque son permisibles valores superiores) indican un buen ajuste relativo, aunque en cualquier caso, un buen valor  $\chi^2$  no garantiza una respuesta correcta en este tipo de experimentos. En el caso aquí descrito, el valor  $\chi^2$  para el ajuste es de 3.21, que pese a no ser un valor ideal, es aún suficientemente bajo. Como puede observarse, existe un leve desajuste entre el modelo atómico obtenido y los datos experimentales, que puede deberse a dos causas: bien a la presencia de agregación en la muestra, o bien a la flexibilidad del complejo. Dado que no se detectaron signos evidentes de agregación ni durante la preparación de la muestra ni la posterior toma de datos, es muy probable que el motivo de este desajuste sea por tanto la elevada flexibilidad del complejo, presentando el complejo una conformación extendida en solución. Es probable que el fragmento entre los residuos 128 y 164 de IP<sub>3</sub>



3-K A se encuentre desordenado (**Figura 4.65b y c**), no presentando una predicción de estructura secundaria definida. En cuanto al conector entre el dominio de unión a CaM y el dominio kinasa, es probable que este se encuentre también desordenado contribuyendo a dicha flexibilidad (**Figura 4.65c**).

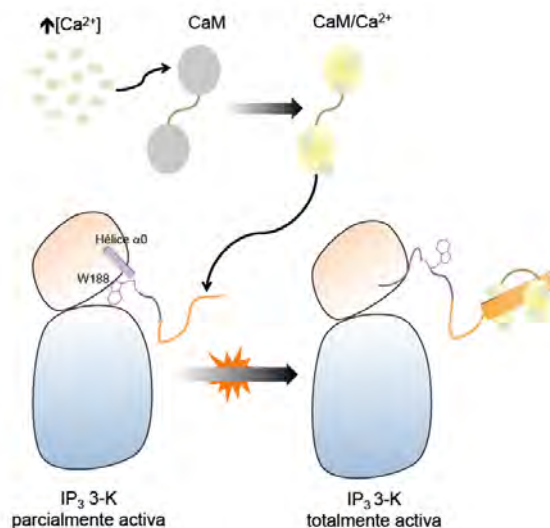


**Figura 4.65. Modelo a alta resolución del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub>.** a) Ajuste entre los datos experimentales y los modelos a baja y alta resolución b) Superposición del modelo a baja resolución y el modelo a alta resolución obtenido mediante ajuste de cuerpo rígido de las estructuras cristalográficas. La IP<sub>3</sub> 3-K<sub>(128-461)</sub> se representa en color rojo, mientras que la CaM se representa en color azul. c) Resultado de la combinación de modelos a alta resolución y baja resolución del complejo CaM/IP<sub>3</sub> 3-K<sub>(128-461)</sub>.

### 4.3.9 Un posible mecanismo de activación de la IP<sub>3</sub> 3-K mediado por CaM

A la vista de los datos obtenidos mediante SAXS (**Figura 4.65**) y la elevada flexibilidad del complejo IP<sub>3</sub> 3-K/CaM es improbable que la CaM active a la IP<sub>3</sub> 3-K mediante una interacción directa. Es más probable que dicha activación se produzca siguiendo un modelo de tipo “pseudosustrato” (Haiech *et al.*, 2011), un modelo que ya había sido propuesto para la activación de la IP<sub>3</sub> 3-K por CaM (González *et al.*, 2004; Chamberlain *et al.*, 2005). En dicho modelo, CaM facilita el desbloqueo del sitio activo de la enzima mediante la unión al CBD. En el caso de la IP<sub>3</sub> 3-K, la unión de la CaM al CBD induce el movimiento de la hélice  $\alpha 0$  de la IP<sub>3</sub> 3-K. Esta  $\alpha 0$  contiene un residuo de triptófano (Trp188) capaz de mimetizar las interacciones del anillo de adenina del ATP en el sitio de unión al nucleótido, el cual actúa por tanto como un dominio de autoinhibición o pseudosustrato. Un mutante del residuo equivalente al Trp188 en la isoforma B humana (W666A) provoca un aumento de la actividad enzimática, lo que apoya el posible rol de este residuo como residuo responsable de la autoinhibición de la

enzima (Chamberlain *et al.*, 2005). Un esquema de este proceso se presenta en la **Figura 4.66**.



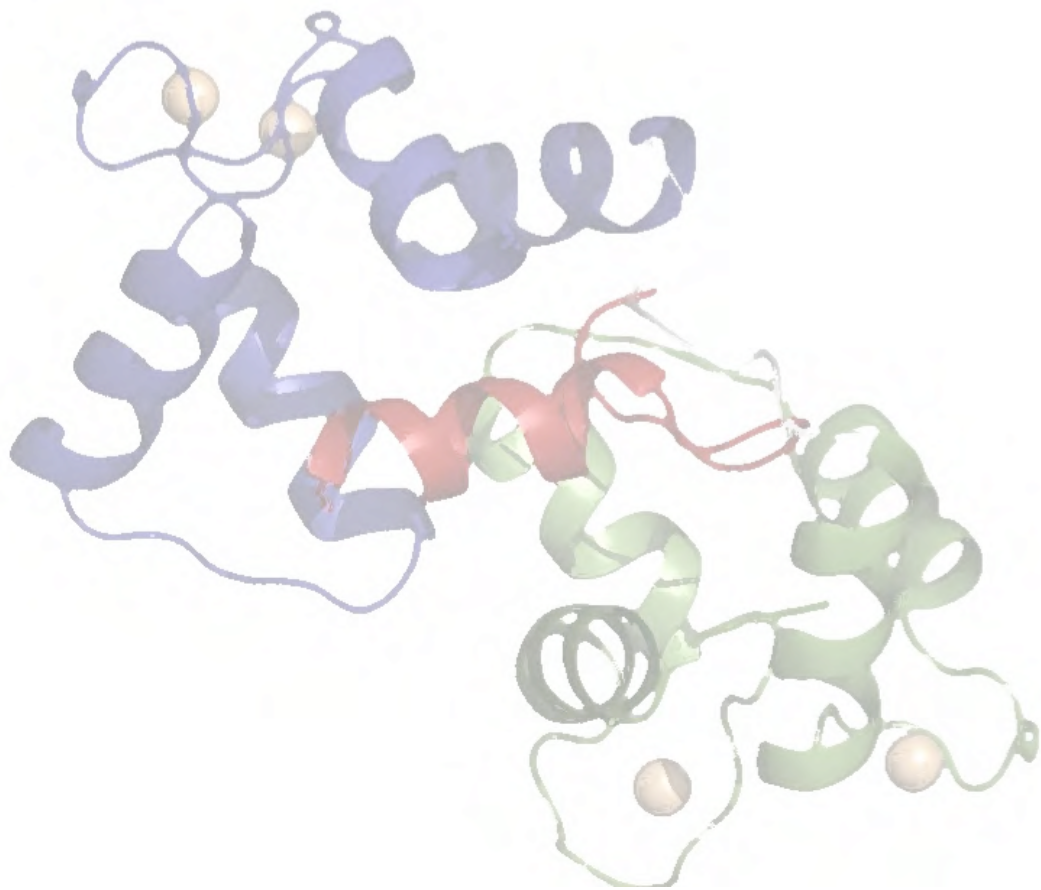
**Figura 4.66. Modelo de la activación de IP<sub>3</sub> 3-K por CaM.** IP<sub>3</sub> 3-K se encuentra parcialmente activa debido a la presencia de W188 (hélice  $\alpha 0$ ) en el sitio de unión a ATP. Al producirse un aumento de  $Ca^{2+}$  en el medio, éste se une a la CaM formando el complejo  $Ca^{2+}/CaM$ . Este complejo es capaz de unirse al sitio de unión a CaM (resaltado en naranja en el esquema). La unión induce la adquisición de una conformación cerrada en la CaM y la formación de una hélice en el CBD, dominio de unión a la CaM, provocando la desaparición de la  $\alpha 0$  y el desbloqueo total del centro activo.

### 4.3.10 Conclusión

Gracias al trabajo aquí expuesto se ha logrado caracterizar por vez primera la unión entre el dominio de unión a CaM de la IP<sub>3</sub> 3-K (CBD) y la CaM. Los datos estructurales obtenidos indican que es muy probable que este dominio de unión pertenezca a una nueva familia de dominios de unión (clase 1-13). Por otro lado, los datos estructurales obtenidos mediante dispersión a bajo ángulo (SAXS) permiten proponer un modelo de activación de IP<sub>3</sub> 3-K por CaM de tipo “pseudosustrato”.







## 5 CONCLUSIONES Y PERSPECTIVAS



## 5.1 Conclusiones

### 5.1.1 IP<sub>5</sub> 2-K

1. Se han caracterizado estructuralmente complejos binarios (con inositido) y ternarios (con inositido y nucleótido) de la IP<sub>5</sub> 2-K de *A. thaliana*. La enzima presenta una estructura  $\alpha+\beta$  compuesta de dos lóbulos N y C-terminal, unidos por un bucle de conexión, mostrando un lóbulo de unión a IP único.
2. El sitio de unión a ATP se encuentra situado en la interfase formada por las láminas  $\beta$  de los dos lóbulos N y C-terminal, estando residuos de ambos lóbulos implicados en la unión del compuesto.
3. El sitio de unión a IP está principalmente conformado por una región  $\alpha$  helicoidal situada en el lóbulo C-terminal, denominada lóbulo C<sub>IP</sub>. La mayoría de residuos fundamentales para la unión de IP<sub>5</sub> y la conversión de éste en IP<sub>6</sub> provienen de esta región, si bien el lóbulo N-terminal también aporta algunos residuos clave.
4. Se han identificado las interacciones del nucleótido y el inositido en el centro activo, determinándose los mecanismos moleculares de reconocimiento de estos compuestos, y el mecanismo enzimático involucrado en la producción del ácido fítico. El rol de los diferentes residuos ha sido confirmado mediante experimentos de mutagénesis dirigida.
5. Se ha caracterizado un sitio de unión a Zn<sup>2+</sup>. Este metal se encuentra probablemente relacionado con el correcto plegamiento de la IP<sub>5</sub> 2-K.
6. La IP<sub>5</sub> 2-K pertenece a la familia de las IPKs, si bien es el miembro más lejano de esta familia. Por tanto, el plegamiento del lóbulo N-terminal y el reconocimiento de ATP se encuentra conservado con la familia de las proteína kinasa.
7. Se ha caracterizado estructuralmente la forma libre de IP<sub>5</sub> 2-K de *A. thaliana*, la cual presenta una forma notablemente abierta en comparación con los complejos cristalizados en presencia de inositido.
8. Se ha caracterizado estructuralmente la forma unida al nucleótido de la IP<sub>5</sub> 2-K de *A. thaliana*, la cual muestra una conformación intermedia entre las formas abierta (libre) y cerrada (unida a inositido).

9. Se ha determinado el papel del lóbulo N-terminal en la unión de sustratos, el cambio conformacional y la discriminación de sustratos enantioméricos.
10. Se ha propuesto un mecanismo mediante distintos ensayos bioquímicos que explica el ciclo catalítico de la IP<sub>5</sub> 2-K.

### 5.1.2 IP<sub>3</sub> 3-K

1. Se ha purificado, cristalizado y determinado la estructura del complejo entre CBD IP<sub>3</sub> 3-K A y la CaM de *H. sapiens*.
2. Se han caracterizado los determinantes de la unión entre estas dos proteínas.
3. Se ha identificado una nueva clase de dominio de unión a CaM caracterizándose por primera vez el motivos de unión 1-13-16.
4. La combinación de datos de alta y baja resolución ha permitido proponer un modelo de activación de la IP<sub>3</sub> 3-K por CaM.

## 5.2 Perspectivas

### 5.2.1 IP<sub>5</sub> 2-K

El desarrollo de este trabajo ha incrementado notablemente el conocimiento a nivel estructural y funcional sobre esta enzima, conociéndose ahora los determinantes que rigen la unión de los ligandos, los residuos que participan en la catálisis y los cambios conformacionales que la enzima sufre durante ésta. De hecho, la IP<sub>5</sub> 2-K es la primera InsP K cuyo ciclo catalítico ha sido totalmente caracterizado mediante cristalografía de rayos X.

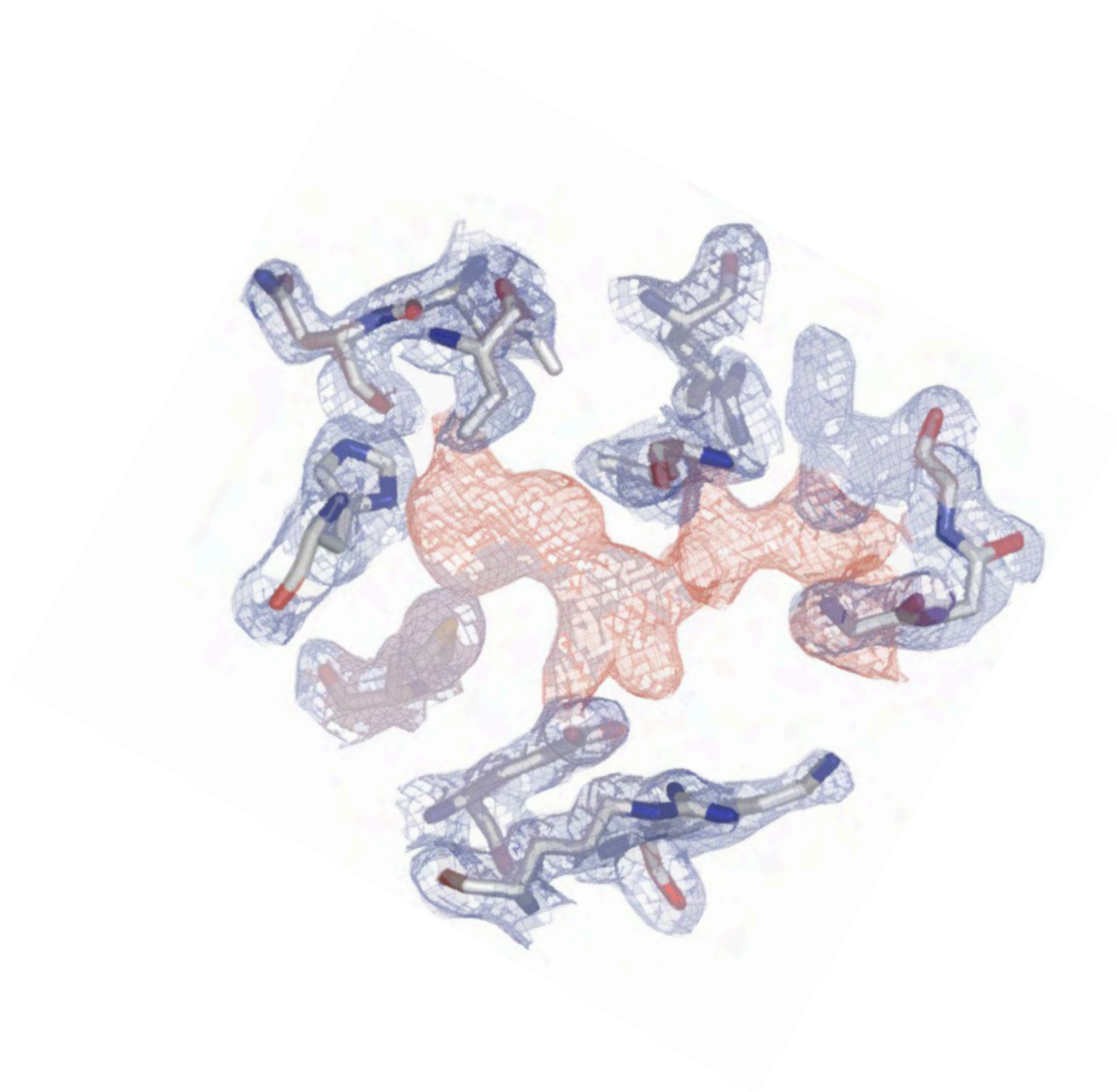
Pese a ser una de las enzimas claves en la ruta de los IPs, no existe ningún inhibidor específico de la IP<sub>5</sub> 2-K. El conocimiento de la estructura y los cambios conformacionales permitirán el desarrollo de inhibidores para la enzima. Estos inhibidores son necesarios como herramienta para seguir ampliando nuestros conocimientos sobre la ruta de los inositol fosfatos, con potenciales aplicaciones clínicas. Por otro lado, el conocimiento de la estructura de la enzima puede ayudar en el desarrollo de especies vegetales cuyas semillas tengan un bajo contenido en fitato (mutantes *lpa*).

### 5.2.2 IP<sub>3</sub> 3-K

Gracias al trabajo aquí presentando se ha logrado caracterizar por primera vez la unión entre el CBD de una IP<sub>3</sub> 3-K y la CaM, identificándose una nueva clase de CBDs (de tipo 1-13-16). Para corroborar esto, se han diseñado una serie de mutantes del CBD: W167A, I170A, R171A, V174A, L176A, I179A y F182A. Estos mutantes serán producidos y la unión entre CaM y los diferentes mutantes de la IP<sub>3</sub> 3-K será monitorizada mediante diferentes técnicas bioquímicas y/o biofísicas.

Por otro lado, y mediante la combinación de los datos obtenidos mediante cristalografía de rayos X y dispersión a bajo ángulo se ha realizado un modelo de la interacción entre la IP<sub>3</sub> 3-K<sub>(128-461)</sub> y la CaM. La obtención de datos de dispersión a bajo ángulo de diferentes tipos de muestras permitirá validar y optimizar este modelo.





## 6 BIBLIOGRAFÍA





- Abdullah M., Hughes, P.J., Craxton, A., Gigg, R., Desai, T., Marecek, J.F., Prestwich, G.D. y Shears, S.B. (1992). Purification and characterization of inositol-1,3,4-trisphosphate 5/6-kinase from rat liver using an inositol hexakisphosphate affinity column. *J Biol Chem.* **267** (31): 22340-22345.
- Altschul S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. y Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25** (17): 3389-3402.
- Angulo I., Acebrón, I., de las Rivas, B., Muñoz, R., Rodríguez-Crespo, I., Menéndez, M., García, P., Tateno, H., Goldstein, I.J., Pérez-Agote, B. y Mancheño, J.M. (2011). High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile beta-trefoil lectin domain from the mushroom *Laetiporus sulphureus*. *Glycobiology.* **21** (10): 1349-1361.
- Ataman Z.A., Gakhar, L., Sorensen, B.R., Hell, J.W. y Shea, M.A. (2007). The NMDA receptor NR1 C1 region bound to calmodulin: structural insights into functional differences between homologous domains. *Structure.* **15** (12): 1603-1617.
- Bang S., Kim, S., Dailey, M.J., Chen, Y., Moran, T.H., Snyder, S.H. y Kim, S.F. (2012). AMP-activated protein kinase is physiologically regulated by inositol polyphosphate multikinase. *Proc Natl Acad Sci U S A.* **109** (2): 616-620.
- Barker C.J., Illies, C., Gaboardi, G.C. y Berggren, P.O. (2009). Inositol pyrophosphates: structure, enzymology and function. *Cell Mol Life Sci.* **66** (24): 3851-3871.
- Bennett M., Onnebo, S.M., Azevedo, C. y Saiardi, A. (2006). Inositol pyrophosphates: metabolism and signaling. *Cell Mol Life Sci.* **63** (5): 552-564.
- Bergfors T. (2003). Seeds to crystals. *J Struct Biol.* **142** (1): 66-76.
- Bergfors T. (2009). Protein Crystallization. International University Line.
- Berman H., Henrick, K. y Nakamura, H. (2003). Announcing the worldwide Protein Data Bank. *Nat Struct Biol.* **10** (12): 980.
- Berridge M.J. (1987). Inositol trisphosphate and diacylglycerol: two interacting second messengers. *Annu Rev Biochem.* **56** 159-193.
- Berridge M.J. (2009). Inositol trisphosphate and calcium signalling mechanisms. *Biochim Biophys Acta.* **1793** (6): 933-940.
- Berridge M.J. y Irvine, R.F. (1984). Inositol trisphosphate, a novel second messenger in cellular signal transduction. *Nature.* **312** (5992): 315-321.
- Bhandari R., Juluri, K.R., Resnick, A.C. y Snyder, S.H. (2008). Gene deletion of inositol hexakisphosphate kinase 1 reveals inositol pyrophosphate regulation of insulin secretion, growth, and spermiogenesis. *Proc Natl Acad Sci U S A.* **105** (7): 2349-2353.
- Blaszczyk J., Shi, G., Yan, H. y Ji, X. (2000). Catalytic center assembly of HPPK as revealed by the crystal structure of a ternary complex at 1.25 Å resolution. *Structure.* **8** (10): 1049-1058.
- Bragg W.H. y Bragg, W.L. (1913). The Reflection of X-rays by Crystals. *Proc. R. Soc. Lond. A.* **88** (605): 439-443.
- Brehm M.A., Schreiber, I., Bertsch, U., Wegner, A. y Mayr, G.W. (2004). Identification of the actin-binding domain of Ins(1,4,5)P<sub>3</sub> 3-kinase isoform B (IP<sub>3</sub>K-B). *Biochem J.* **382** (Pt 1): 353-362.

- Brehm M.A., Wundenberg, T., Williams, J., Mayr, G.W. y Shears, S.B.** (2012). A non-catalytic role for inositol 1,3,4,5,6-pentakisphosphate 2-kinase in the synthesis of ribosomal RNA. *J Cell Sci.* 126(2):437-44
- Brunger A.T.** (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature.* **355** (6359): 472-475.
- Caddick S.E., Harrison, C.J., Stavridou, I., Johnson, S. y Brearley, C.A.** (2007). A lysine accumulation phenotype of ScIpk2Delta mutant yeast is rescued by *Solanum tuberosum* inositol phosphate multikinase. *Biochem J.* **403** (3): 381-389.
- Carvin D., Islam, S.A., Sternberg, M.J., Blundell, T.L., Stubbs, M.T. y Huber, R.** (2006). Isomorphous replacement. En *International Tables of Crystallography. F: Crystallography of biological macromolecules.* (12) 247-255.
- Cebrian D., Tapia, A., Real, A. y Morcillo, M.A.** (2007). Inositol hexaphosphate: a potential chelating agent for uranium. *Radiat Prot Dosimetry.* **127** (1-4): 477-479.
- Chakraborty A., Koldobskiy, M.A., Bello, N.T., Maxwell, M., Potter, J.J., Juluri, K.R., Maag, D., Kim, S., Huang, A.S., Dailey, M.J., Saleh, M., Snowman, A.M., Moran, T.H., Mezey, E. y Snyder, S.H.** (2010). Inositol pyrophosphates inhibit Akt signaling, thereby regulating insulin sensitivity and weight gain. *Cell.* **143** (6): 897-910.
- Chakraborty A., Koldobskiy, M.A., Sixt, K.M., Juluri, K.R., Mustafa, A.K., Snowman, A.M., van Rossum, D.B., Patterson, R.L. y Snyder, S.H.** (2008). HSP90 regulates cell survival via inositol hexakisphosphate kinase-2. *Proc Natl Acad Sci U S A.* **105** (4): 1134-1139.
- Chamberlain P.P., Qian, X., Stiles, A.R., Cho, J., Jones, D.H., Lesley, S.A., Grabau, E.A., Shears, S.B. y Spraggon, G.** (2007). Integration of inositol phosphate signaling pathways via human ITPK1. *J Biol Chem.* **282** (38): 28117-28125.
- Chamberlain P.P., Sandberg, M.L., Sauer, K., Cooke, M.P., Lesley, S.A. y Spraggon, G.** (2005). Structural insights into enzyme regulation for inositol 1,4,5-trisphosphate 3-kinase B. *Biochemistry.* **44** (44): 14486-14493.
- Chayen N.E. y Saridakis, E.** (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods.* **5** (2): 147-153.
- Cheek S., Ginalski, K., Zhang, H. y Grishin, N.V.** (2005). A comprehensive update of the sequence and structure classification of kinases. *BMC Struct Biol.* **5** 6.
- Cheek S., Zhang, H. y Grishin, N.V.** (2002). Sequence and structure classification of kinases. *J Mol Biol.* **320** (4): 855-881.
- Chen V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. y Richardson, D.C.** (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr.* **66** (Pt 1): 12-21.
- Cheung J.C., Salerno, B. y Hanakahi, L.A.** (2008). Evidence for an inositol hexakisphosphate-dependent role for Ku in mammalian nonhomologous end joining that is independent of its role in the DNA-dependent protein kinase. *Nucleic Acids Res.* **36** (17): 5713-5726.
- Chin D. y Means, A.R.** (2000). Calmodulin: a prototypical calcium sensor. *Trends Cell Biol.* **10** (8): 322-328.

- Choi K.Y., Kim, H.K., Lee, S.Y., Moon, K.H., Sim, S.S., Kim, J.W., Chung, H.K. y Rhee, S.G.** (1990). Molecular cloning and expression of a complementary DNA for inositol 1,4,5-trisphosphate 3-kinase. *Science*. **248** (4951): 64-66.
- Chu D.M., Corbin, J.D., Grimes, K.A. y Francis, S.H.** (1997). Activation by cyclic GMP binding causes an apparent conformational change in cGMP-dependent protein kinase. *J Biol Chem*. **272** (50): 31922-31928.
- Communi D., Dewaste, V. y Erneux, C.** (1999). Calcium-calmodulin-dependent protein kinase II and protein kinase C-mediated phosphorylation and activation of D-myo-inositol 1,4, 5-trisphosphate 3-kinase B in astrocytes. *J Biol Chem*. **274** (21): 14734-14742.
- Communi D., Vanweyenberg, V. y Erneux, C.** (1997). D-myo-inositol 1,4,5-trisphosphate 3-kinase A is activated by receptor activation through a calcium : calmodulin-dependent protein kinase II phosphorylation mechanism. *EMBO J*. **16** (8): 1943-1952.
- Cowtan K.** (2008). Fitting molecular fragments into electron density. *Acta Crystallogr D Biol Crystallogr*. **64** (Pt 1): 83-89.
- Cowtan K.** (2010). Recent developments in classical density modification. *Acta Crystallogr D Biol Crystallogr*. **66** (Pt 4): 470-478.
- Cox S., Radzio-Andzelm, E. y Taylor, S.S.** (1994). Domain movements in protein kinases. *Curr Opin Struct Biol*. **4** (6): 893-901.
- Dauter Z. y Wilson, K.S.** (2006). Monochromatic data collection. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules*. **9.1**. 177-195.
- Dewaste V., Moreau, C., De Smedt, F., Bex, F., De Smedt, H., Wuytack, F., Missiaen, L. y Erneux, C.** (2003). The three isoenzymes of human inositol-1,4,5-trisphosphate 3-kinase show specific intracellular localization but comparable  $\text{Ca}^{2+}$  responses on transfection in COS-7 cells. *Biochem J*. **374** (Pt 1): 41-49.
- Dewaste V., Pouillon, V., Moreau, C., Shears, S., Takazawa, K. y Erneux, C.** (2000). Cloning and expression of a cDNA encoding human inositol 1,4,5-trisphosphate 3-kinase C. *Biochem J*. **352** Pt 2 343-351.
- Drum C.L., Yan, S.Z., Bard, J., Shen, Y.Q., Lu, D., Soelaiman, S., Grabarek, Z., Bohm, A. y Tang, W.J.** (2002). Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. *Nature*. **415** (6870): 396-402.
- Eisenberg D., Schwarz, E., Komaromy, M. y Wall, R.** (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol*. **179** (1): 125-142.
- Emsley P., Lohkamp, B., Scott, W.G. y Cowtan, K.** (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. **66** (Pt 4): 486-501.
- Endo-Streeter S., Tsui, M.K., Odom, A.R., Block, J. y York, J.D.** (2012). Structural studies and protein engineering of inositol phosphate multikinase. *J Biol Chem*. **287** (42): 35360-35369.
- Engh R.A. y Huber, R.** (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr A*. **47** (Pt 4): 392-400.
- Erneux C., Moreau, C., Vandermeers, A. y Takazawa, K.** (1993). Interaction of calmodulin with a putative calmodulin-binding domain of inositol 1,4,5-trisphosphate 3-kinase. Effects of synthetic peptides and site-directed mutagenesis of Trp165. *Eur J Biochem*. **214** (2): 497-501.

- Evans P.** (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr.* **62** (Pt 1): 72-82.
- Evans P.R.** (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr.* **67** (Pt 4): 282-292.
- Franke D. y Svergun, D.I.** (2009). DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **42** 342-346.
- French G.S. y Wilson, K.S.** (1978). On the treatment of negative intensity observations. *Acta Crystallogr A.* **34** 517-525.
- Garman E.F. y Schneider, T.R.** (1997). Macromolecular Cryocrystallography. *J. Appl. Cryst.* **30** 211-237.
- Giacovazzo C.** (2011). Fundamentals of Crystallography. Oxford University Press.
- Giegé R. y McPherson, A.** (2006). Crystallization. En *International Tables of Crystallography. Volumen F: Crystallography of Biological Macromolecules.* (4.1). 81-93.
- Gifford J.L., Ishida, H. y Vogel, H.J.** (2011). Fast methionine-based solution structure determination of calcium-calmodulin complexes. *J Biomol NMR.* **50** (1): 71-81.
- González B., Schell, M.J., Letcher, A.J., Veprintsev, D.B., Irvine, R.F. y Williams, R.L.** (2004). Structure of a human inositol 1,4,5-trisphosphate 3-kinase: substrate binding reveals why it is not a phosphoinositide 3-kinase. *Mol Cell.* **15** (5): 689-701.
- Gosein V., Leung, T.F., Krajden, O. y Miller, G.J.** (2012). Inositol phosphate-induced stabilization of inositol 1,3,4,5,6-pentakisphosphate 2-kinase and its role in substrate specificity. *Protein Sci.* **21** (5): 737-742.
- Green D.W., Ingram, V.M. y Perutz, M.F.** (1954). The Structure of Haemoglobin. IV. Sign Determination by the Isomorphous Replacement Method. *Proc. R. Soc. Lond. A.* **225** (1162): 287-307.
- Gruner S.M., Eikenberry, E.F. y Tate, M.W.** (2006). Comparision of X-ray detectors. En *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules.* 7.1. 143-147.
- Haiech J., Audran, E., Feve, M., Ranjeva, R. y Kilhoffer, M.C.** (2011). Revisiting intracellular calcium signaling semantics. *Biochimie.* **93** (12): 2029-2037.
- Hatch A.J. y York, J.D.** (2010). SnapShot: Inositol phosphates. *Cell.* **143** (6): 1030-1030 e1031.
- Hayashi N., Matsubara, M., Takasaki, A., Titani, K. y Taniguchi, H.** (1998). An expression system of rat calmodulin using T7 phage promoter in *Escherichia coli*. *Protein Expr Purif.* **12** (1): 25-28.
- Hayward S. y Berendsen, H.J.** (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins.* **30** (2): 144-154.
- Helliwell J.R.** (2006). Synchrotron-radiation instrumentation, methods and scientific utilization. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules.* 8.1. 155-166.
- Hendrickson W.A., Horton, J.R. y LeMaster, D.M.** (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9** (5): 1665-1672.

- Hendrickson W.A., Smith, J.L. y Sheriff, S.** (1985). Direct Phase determination based on anomalous scattering. *Methods Enzymol.* **115** 41-55.
- Hendrickson W.A. y Teeter, M.M.** (1981). Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature.* **290** 107-113.
- Hermosura M.C., Takeuchi, H., Fleig, A., Riley, A.M., Potter, B.V., Hirata, M. y Penner, R.** (2000). InsP4 facilitates store-operated calcium influx by inhibition of InsP3 5-phosphatase. *Nature.* **408** (6813): 735-740.
- Ho M.W., Yang, X., Carew, M.A., Zhang, T., Hua, L., Kwon, Y.U., Chung, S.K., Adelt, S., Vogel, G., Riley, A.M., Potter, B.V. y Shears, S.B.** (2002). Regulation of Ins(3,4,5,6)P(4) signaling by a reversible kinase/phosphatase. *Curr Biol.* **12** (6): 477-482.
- Holm L. y Rosenstrom, P.** (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38** (Web Server issue): W545-549.
- Holmes W. y Jogl, G.** (2006). Crystal structure of inositol phosphate multikinase 2 and implications for substrate specificity. *J Biol Chem.* **281** (49): 38109-38116.
- Hughes S.H. y Stock, A.M.** (2006). Preparing recombinant proteins for X-ray crystallography. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules.* **3.1.** 65-80.
- Hurrell R.F.** (2003). Influence of vegetable protein sources on trace element and mineral bioavailability. *J Nutr.* **133** (9): 2973S-2977S.
- Ikura M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. y Bax, A.** (1992). Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science.* **256** (5057): 632-638.
- Irvine R.** (2001). Inositol phosphates: Does IP(4) run a protection racket? *Curr Biol.* **11** (5): R172-174.
- Irvine R.F., Letcher, A.J., Heslop, J.P. y Berridge, M.J.** (1986). The inositol tris/tetrakisphosphate pathway--demonstration of Ins(1,4,5)P3 3-kinase activity in animal tissues. *Nature.* **320** (6063): 631-634.
- Irvine R.F. y Schell, M.J.** (2001). Back in the water: the return of the inositol phosphates. *Nat Rev Mol Cell Biol.* **2** (5): 327-338.
- Ives E.B., Nichols, J., Wentz, S.R. y York, J.D.** (2000). Biochemical and functional characterization of inositol 1,3,4,5, 6-pentakisphosphate 2-kinases. *J Biol Chem.* **275** (47): 36575-36583.
- Jamison R.S., Newcomer, M.E. y Ong, D.E.** (1998). Detection of conformational changes in cellular retinoid-binding proteins by limited proteolysis. *Methods Mol Biol.* **89** 165-176.
- Jancarik J. y Kim, S.H.** (1991). Sparse matrix sampling: A screening method for the crystallization of macromolecules. *J. Appl. Cryst.* **24** 401-409.
- Janmey P.A.** (1994). Phosphoinositides and calcium as regulators of cellular actin assembly and disassembly. *Annu Rev Physiol.* **56** 169-191.
- Johnson H.W. y Schell, M.J.** (2009). Neuronal IP<sub>3</sub> 3-kinase is an F-actin-bundling protein: role in dendritic targeting and regulation of spine morphology. *Mol Biol Cell.* **20** (24): 5166-5180.
- Johnson L.F. y Tate, M.E.** (1969). The structure of myo-inositol pentaphosphates. *Ann N Y Acad Sci.* **165** (2): 526-532.



- Jones T.A., Zou, J.Y., Cowan, S.W. y Kjeldgaard, M.** (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A*. **47** ( Pt 2) 110-119.
- Kabsch W.** (2006). Integration, scaling, space-group assignment and post refinement. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules*. 11.3. 218-225.
- Kahn R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J.L., Dideberg, O. y Wery, J.P.** (1985). Crystal structure study of *Opsanus tau* parvalbumin by multiwavelength anomalous diffraction. *FEBS Lett.* **179** (1): 133-137.
- Kim S., Kim, S.F., Maag, D., Maxwell, M.J., Resnick, A.C., Juluri, K.R., Chakraborty, A., Koldobskiy, M.A., Cha, S.H., Barrow, R., Snowman, A.M. y Snyder, S.H.** (2011). Amino acid signaling to mTOR mediated by inositol polyphosphate multikinase. *Cell Metab.* **13** (2): 215-221.
- Kleywegt G.J.** (2000). Validation of protein crystal structures. *Acta Crystallogr D Biol Crystallogr.* **56** (Pt 3): 249-265.
- Koldobskiy M.A., Chakraborty, A., Werner, J.K., Jr., Snowman, A.M., Juluri, K.R., Vandiver, M.S., Kim, S., Heletz, S. y Snyder, S.H.** (2010). p53-mediated apoptosis requires inositol hexakisphosphate kinase-2. *Proc Natl Acad Sci U S A*. **107** (49): 20947-20951.
- Koldobskiy M.A. y Snyder, S.H.** (2011). Inositol pyrophosphates in cell death and life. *Cell Cycle.* **10** (4): 568-570.
- Komander D., Fairservice, A., Deak, M., Kular, G.S., Prescott, A.R., Peter Downes, C., Safrany, S.T., Alessi, D.R. y van Aalten, D.M.** (2004). Structural insights into the regulation of PDK1 by phosphoinositides and inositol phosphates. *EMBO J.* **23** (20): 3918-3928.
- Konarev P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J. y Svergun, D.I.** (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst.* **36** 1277-1282.
- Kovacs E., Harmat, V., Toth, J., Vertessy, B.G., Modos, K., Kardos, J. y Liliom, K.** (2010). Structure and mechanism of calmodulin binding to a signaling sphingolipid reveal new aspects of lipid-protein interactions. *FASEB J.* **24** (10): 3829-3839.
- Kraft P., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E.F., Henrich, B., Johnson, I., Mozzanica, A., Schleputz, C.M., Willmott, P.R. y Schmitt, B.** (2009). Performance of single-photon-counting PILATUS detector modules. *J Synchrotron Radiat.* **16** (Pt 3): 368-375.
- Kranz J.K., Lee, E.K., Nairn, A.C. y Wand, A.J.** (2002). A direct test of the reductionist approach to structural studies of calmodulin activity: relevance of peptide models of target proteins. *J Biol Chem.* **277** (19): 16351-16354.
- Kretsinger R.H. y Nockolds, C.E.** (1973). Carp muscle calcium-binding protein. II. Structure determination and general description. *J Biol Chem.* **248** (9): 3313-3326.
- Kuboniwa H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B. y Bax, A.** (1995). Solution structure of calcium-free calmodulin. *Nat Struct Biol.* **2** (9): 768-776.
- Kucharski R. y Maleszka, R.** (2002). Molecular profiling of behavioural development: differential expression of mRNAs for inositol 1,4,5-trisphosphate 3-kinase isoforms in naive and experienced honeybees (*Apis mellifera*). *Brain Res Mol Brain Res.* **99** (2): 92-101.



- Laemmli U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227** (5259): 680-685.
- Larkin M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. y Higgins, D.G.** (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*. **23** (21): 2947-2948.
- Laskowski R.A., MacArthur, M.W., Moss, D.S. y Thornton, J.M.** (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26** (Pt 2): 283-291.
- Lee B. y Richards, F.M.** (1971). The interpretation of protein structures: estimation of static accessibility. *J Mol Biol.* **55** (3): 379-400.
- Lee Y.S., Mulugu, S., York, J.D. y O'Shea, E.K.** (2007). Regulation of a cyclin-CDK-CDK inhibitor complex by inositol pyrophosphates. *Science*. **316** (5821): 109-112.
- Lemmon M.A.** (2003). Phosphoinositide recognition domains. *Traffic*. **4** (4): 201-213.
- Leslie A.G.** (2006). The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr.* **62** (Pt 1): 48-57.
- Leslie A.G.** (2006). Integration of macromolecular diffraction data. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules*. **11.1**. 212-217.
- Liscovitch M. y Cantley, L.C.** (1995). Signal transduction and membrane traffic: the P1TP/phosphoinositide connection. *Cell*. **81** (5): 659-662.
- Lloyd-Burton S.M., Yu, J.C., Irvine, R.F. y Schell, M.J.** (2007). Regulation of inositol 1,4,5-trisphosphate 3-kinases by calcium and localization in cells. *J Biol Chem*. **282** (13): 9526-9535.
- Long F., Vagin, A.A., Young, P. y Murshudov, G.N.** (2008). BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr.* **64** (Pt 1): 125-132.
- Lu P.J., Shieh, W.R. y Chen, C.S.** (1996). Antagonistic effect of inositol pentakisphosphate on inositol triphosphate receptors. *Biochem Biophys Res Commun*. **220** (3): 637-642.
- Luo H.R., Huang, Y.E., Chen, J.C., Saiardi, A., Iijima, M., Ye, K., Huang, Y., Nagata, E., Devreotes, P. y Snyder, S.H.** (2003). Inositol pyrophosphates mediate chemotaxis in *Dictyostelium* via pleckstrin homology domain-PtdIns(3,4,5)P<sub>3</sub> interactions. *Cell*. **114** (5): 559-572.
- Luo H.R., Saiardi, A., Nagata, E., Ye, K., Yu, H., Jung, T.S., Luo, X., Jain, S., Sawa, A. y Snyder, S.H.** (2001). GRAB: a physiologic guanine nucleotide exchange factor for Rab3A, which interacts with inositol hexakisphosphate kinase. *Neuron*. **31** (3): 439-451.
- Luo H.R., Saiardi, A., Yu, H., Nagata, E., Ye, K. y Snyder, S.H.** (2002). Inositol pyrophosphates are required for DNA hyperrecombination in protein kinase c1 mutant yeast. *Biochemistry*. **41** (8): 2509-2515.
- Macbeth M.R., Schubert, H.L., Vandemark, A.P., Lingam, A.T., Hill, C.P. y Bass, B.L.** (2005). Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science*. **309** (5740): 1534-1539.
- Maffucci T., Piccolo, E., Cumashi, A., Iezzi, M., Riley, A.M., Saiardi, A., Godage, H.Y., Rossi, C., Broggin, M., Iacobelli, S., Potter, B.V., Innocenti, P. y Falasca, M.** (2005). Inhibition of the phosphatidylinositol 3-kinase/Akt pathway

- by inositol pentakisphosphate results in antiangiogenic and antitumor effects. *Cancer Res.* **65** (18): 8339-8349.
- Matthews B.W.** (1968). Solvent content of protein crystals. *J Mol Biol.* **33** (2): 491-497.
- McCoy A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. y Read, R.J.** (2007). Phaser crystallographic software. *J Appl Crystallogr.* **40** (Pt 4): 658-674.
- McPherson A.** (1999). Crystallization of Biological Macromolecules. Cold Spring Harbor Laboratory Press.
- McPherson A.** (2004). Introduction to protein crystallization. *Methods.* **34** (3): 254-265.
- McPherson A.** (2009). Introduction to Macromolecular Crystallography. John Wiley & Sons, Inc.
- Meador W.E., Means, A.R. y Quioco, F.A.** (1992). Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science.* **257** (5074): 1251-1255.
- Meador W.E., Means, A.R. y Quioco, F.A.** (1993). Modulation of calmodulin plasticity in molecular recognition on the basis of X ray structures. *Science.* **262** (5140): 1718-1721.
- Means A.R., Tash, J.S. y Chafouleas, J.G.** (1982). Physiological implications of the presence, distribution, and regulation of calmodulin in eukaryotic cells. *Physiol Rev.* **62** (1): 1-39.
- Meselson M. y Yuan, R.** (1968). DNA restriction enzyme from *E. coli*. *Nature.* **217** (5134): 1110-1114.
- Messenguy F., Dubois, E. y Boonchird, C.** (1991). Determination of the DNA-binding sequences of ARGR proteins to arginine anabolic and catabolic promoters. *Mol Cell Biol.* **11** (5): 2852-2863.
- Mikoshiba K.** (2007). IP<sub>3</sub> receptor/Ca<sup>2+</sup> channel: from discovery to new signaling concepts. *J Neurochem.* **102** (5): 1426-1446.
- Miller G.J. y Hurley, J.H.** (2004). Crystal structure of the catalytic core of inositol 1,4,5-trisphosphate 3-kinase. *Mol Cell.* **15** (5): 703-711.
- Miller G.J., Wilson, M.P., Majerus, P.W. y Hurley, J.H.** (2005). Specificity determinants in inositol polyphosphate synthesis: crystal structure of inositol 1,3,4-trisphosphate 5/6-kinase. *Mol Cell.* **18** (2): 201-212.
- Miroux B. y Walker, J.E.** (1996). Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol.* **260** (3): 289-298.
- Moffatt B.A. y Studier, F.W.** (1987). T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell.* **49** (2): 221-227.
- Montpetit B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M. y Weis, K.** (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP<sub>6</sub> in mRNA export. *Nature.* **472** (7342): 238-242.
- Morrison B.H., Bauer, J.A., Kalvakolanu, D.V. y Lindner, D.J.** (2001). Inositol hexakisphosphate kinase 2 mediates growth suppressive and apoptotic effects of interferon-beta in ovarian carcinoma cells. *J Biol Chem.* **276** (27): 24965-24970.
- Morrison B.H., Bauer, J.A., Lupica, J.A., Tang, Z., Schmidt, H., DiDonato, J.A. y Lindner, D.J.** (2007). Effect of inositol hexakisphosphate kinase 2 on

- transforming growth factor beta-activated kinase 1 and NF-kappaB activation. *J Biol Chem.* **282** (21): 15349-15356.
- Morrison B.H., Haney, R., Lamarre, E., Drazba, J., Prestwich, G.D. y Lindner, D.J.** (2009). Gene deletion of inositol hexakisphosphate kinase 2 predisposes to aerodigestive tract carcinoma. *Oncogene.* **28** (25): 2383-2392.
- Mueller-Planitz F., Klinker, H., Ludwigsen, J. y Becker, P.B.** (2013). The ATPase domain of ISWI is an autonomous nucleosome remodeling machine. *Nat Struct Mol Biol.* **20** (1): 82-89.
- Mulugu S., Bai, W., Fridy, P.C., Bastidas, R.J., Otto, J.C., Dollins, D.E., Haystead, T.A., Ribeiro, A.A. y York, J.D.** (2007). A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science.* **316** (5821): 106-109.
- Murphy A.M., Otto, B., Brearley, C.A., Carr, J.P. y Hanke, D.E.** (2008). A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *Plant J.* **56** (4): 638-652.
- Murshudov G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F. y Vagin, A.A.** (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* **67** (Pt 4): 355-367.
- Nagata E., Luo, H.R., Saiardi, A., Bae, B.I., Suzuki, N. y Snyder, S.H.** (2005). Inositol hexakisphosphate kinase-2, a physiologic mediator of cell death. *J Biol Chem.* **280** (2): 1634-1640.
- Nalaskowski M.M., Bertsch, U., Fanick, W., Stockebrand, M.C., Schmale, H. y Mayr, G.W.** (2003). Rat inositol 1,4,5-trisphosphate 3-kinase C is enzymatically specialized for basal cellular inositol trisphosphate phosphorylation and shuttles actively between nucleus and cytoplasm. *J Biol Chem.* **278** (22): 19765-19776.
- Nalaskowski M.M., Deschermeier, C., Fanick, W. y Mayr, G.W.** (2002). The human homologue of yeast ArgRIII protein is an inositol phosphate multikinase with predominantly nuclear localization. *Biochem J.* **366** (Pt 2): 549-556.
- O'Neil K.T. y DeGrado, W.F.** (1990). How calmodulin binds its targets: sequence independent recognition of amphiphilic alpha-helices. *Trends Biochem Sci.* **15** (2): 59-64.
- Odom A.R., Stahlberg, A., Wente, S.R. y York, J.D.** (2000). A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control. *Science.* **287** (5460): 2026-2029.
- Ono Y. y Sorimachi, H.** (2012). Calpains: an elaborate proteolytic system. *Biochim Biophys Acta.* **1824** (1): 224-236.
- Osawa M., Tokumitsu, H., Swindells, M.B., Kurihara, H., Orita, M., Shibamura, T., Furuya, T. y Ikura, M.** (1999). A novel target recognition revealed by calmodulin in complex with Ca<sup>2+</sup> calmodulin-dependent kinase kinase. *Nat Struct Biol.* **6** (9): 819-824.
- Painter J. y Merritt, E.A.** (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr.* **62** (Pt 4): 439-450.
- Papworth C., Bauer, J.C., Braman, J. y Wright, D.A.** (1996). Site-directed mutagenesis in one day with >80% efficiency. *Strategies* **9** (3): 3-4.
- Pattni K. y Banting, G.** (2004). Ins(1,4,5)P3 metabolism and the family of IP3-3Kinases. *Cell Signal.* **16** (6): 643-654.

- Pattni K., Millard, T.H. y Banting, G.** (2003). Calpain cleavage of the B isoform of Ins(1,4,5)P<sub>3</sub> 3-kinase separates the catalytic domain from the membrane anchoring domain. *Biochem J.* **375** (Pt 3): 643-651.
- Payraastre B., Missy, K., Giuriato, S., Bodin, S., Plantavid, M. y Gratacap, M.** (2001). Phosphoinositides: key players in cell signalling, in time and space. *Cell Signal.* **13** (6): 377-387.
- Petoukhov M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D.T., Konarev, P.V. y Svergun, D.I.** (2012). New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Cryst.* **45**, 342-350.
- Petoukhov M.V. y Svergun, D.I.** (2005). Global rigid body modelling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89** 1237-1250.
- Pettersen E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. y Ferrin, T.E.** (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* **25** (13): 1605-1612.
- Pflugrath J.W.** (2004). Macromolecular cryocrystallography—methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods.* **34** (3): 415-423.
- Phillippy B.Q., Ullah, A.H. y Ehrlich, K.C.** (1994). Purification and some properties of inositol 1,3,4,5,6-Pentakisphosphate 2-kinase from immature soybean seeds. *J Biol Chem.* **269** (45): 28393-28399.
- Posternak S.** (1919). Sur la synthèse de l'eter hexaphosphorique de l'inosite avec le principe phosphoorganique de réserve des plantes vertes. *Comp Rend Acad Sci.* **169** 138-140.
- Punta M., Coggill, P.C., Eberhardt, R.Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E.L., Eddy, S.R., Bateman, A. y Finn, R.D.** (2012). The Pfam protein families database. *Nucleic Acids Res.* **40** (Database issue): D290-301.
- Raboy V.** (2001). Seeds for a better future: 'low phytate' grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci.* **6** (10): 458-462.
- Raboy V. y Bowen, D.** (2006). Genetics of inositol polyphosphates. *Subcell Biochem.* **39** 71-101.
- Ramachandran G.N., Ramakrishnan, C. y Sasisekharan, V.** (1963). Stereochemistry of polypeptide chain configurations. *J Mol Biol.* **7** 95-99.
- Rhoads A.R. y Friedberg, F.** (1997). Sequence motifs for calmodulin recognition. *FASEB J.* **11** (5): 331-340.
- Rhodes G.** (2006). Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models. Academic Press.
- Rossmann M.G.** (2006). Automatic indexing of data processing. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules.* **11.** 209-211.
- Rossmann M.G. y Blow, D.M.** (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr* **15** (Pt 1): 24-31.
- Rowan A.S., Nicely, N.I., Cochrane, N., Wlassoff, W.A., Claiborne, A. y Hamilton, C.J.** (2009). Nucleoside triphosphate mimicry: a sugar triazolyl nucleoside as an ATP-competitive inhibitor of *B. anthracis* pantothenate kinase. *Org Biomol Chem.* **7** (19): 4029-4036.

- Rual J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth, F.P. y Vidal, M. (2005). Towards a proteome-scale map of the human protein-protein interaction network. *Nature*. **437** (7062): 1173-1178.
- Rupp B. (2009). Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology. Garland Science.
- Saiardi A. (2012). Cell signalling by inositol pyrophosphates. *Subcell Biochem*. **59** 413-443.
- Saiardi A., Bhandari, R., Resnick, A.C., Snowman, A.M. y Snyder, S.H. (2004). Phosphorylation of proteins by inositol pyrophosphates. *Science*. **306** (5704): 2101-2105.
- Saiardi A., Caffrey, J.J., Snyder, S.H. y Shears, S.B. (2000). Inositol polyphosphate multikinase (ArgR11) determines nuclear mRNA export in *Saccharomyces cerevisiae*. *FEBS Lett*. **468** (1): 28-32.
- Saiardi A., Erdjument-Bromage, H., Snowman, A.M., Tempst, P. y Snyder, S.H. (1999). Synthesis of diphosphoinositol pentakisphosphate by a newly identified family of higher inositol polyphosphate kinases. *Curr Biol*. **9** (22): 1323-1326.
- Saiardi A., Nagata, E., Luo, H.R., Snowman, A.M. y Snyder, S.H. (2001). Identification and characterization of a novel inositol hexakisphosphate kinase. *J Biol Chem*. **276** (42): 39179-39185.
- Saiardi A., Resnick, A.C., Snowman, A.M., Wendland, B. y Snyder, S.H. (2005). Inositol pyrophosphates regulate cell death and telomere length through phosphoinositide 3-kinase-related protein kinases. *Proc Natl Acad Sci U S A*. **102** (6): 1911-1914.
- Sambrook J. y Russell, D.W. (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press.
- Sarmah B., Latimer, A.J., Appel, B. y Went, S.R. (2005). Inositol polyphosphates regulate zebrafish left-right asymmetry. *Dev Cell*. **9** (1): 133-145.
- Schell M.J. (2010). Inositol trisphosphate 3-kinases: focus on immune and neuronal signaling. *Cell Mol Life Sci*. **67** (11): 1755-1778.
- Schell M.J., Erneux, C. y Irvine, R.F. (2001). Inositol 1,4,5-trisphosphate 3-kinase A associates with F-actin and dendritic spines via its N terminus. *J Biol Chem*. **276** (40): 37537-37546.
- Schell M.J., Letcher, A.J., Brearley, C.A., Biber, J., Murer, H. y Irvine, R.F. (1999). PiUS (Pi uptake stimulator) is an inositol hexakisphosphate kinase. *FEBS Lett*. **461** (3): 169-172.
- Scherer J. (1850). Ueber eine neue, aus dem Muskelfleische gewonnene Zuckerart. *Liebigs Ann Chem*. **73** 322-328.
- Schrödinger L. (2010). The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.
- Shears S.B. (2004). How versatile are inositol phosphate kinases? *Biochem J*. **377** (Pt 2): 265-280.



- Shears S.B., Ganapathi, S.B., Gokhale, N.A., Schenk, T.M., Wang, H., Weaver, J.D., Zaremba, A. y Zhou, Y. (2012). Defining signal transduction by inositol phosphates. *Subcell Biochem.* **59** 389-412.
- Sheldrick G.M. (2008). A short history of SHELX. *Acta Crystallogr A.* **64** (Pt 1): 112-122.
- Shen X., Xiao, H., Ranallo, R., Wu, W.H. y Wu, C. (2003). Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science.* **299** (5603): 112-114.
- Shi Y., Azab, A.N., Thompson, M.N. y Greenberg, M.L. (2006). Inositol phosphates and phosphoinositides in health and disease. *Subcell Biochem.* **39** 265-292.
- Sim S.S., Kim, J.W. y Rhee, S.G. (1990). Regulation of D-myo-inositol 1,4,5-trisphosphate 3-kinase by cAMP-dependent protein kinase and protein kinase C. *J Biol Chem.* **265** (18): 10367-10372.
- Sorensen T.L., McAuley, K.E., Flaig, R. y Duke, E.M. (2006). New light for science: synchrotron radiation in structural medicine. *Trends Biotechnol.* **24** (11): 500-508.
- Stephens L.R., Hawkins, P.T., Stanley, A.F., Moore, T., Poyner, D.R., Morris, P.J., Hanley, M.R., Kay, R.R. y Irvine, R.F. (1991). myo-inositol pentakisphosphates. Structure, biological occurrence and phosphorylation to myo-inositol hexakisphosphate. *Biochem J.* **275** ( Pt 2) 485-499.
- Stevens R.C. (2000). High-throughput protein crystallization. *Curr Opin Struct Biol.* **10** (5): 558-563.
- Studier F.W. y Moffatt, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol.* **189** (1): 113-130.
- Stura E.A. y Wilson, I.A. (1992). Seeding techniques. *Crystallization of nucleic acid and proteins. A practical approach.* **5**. 99-125.
- Sun Y., Thompson, M., Lin, G., Butler, H., Gao, Z., Thornburgh, S., Yau, K., Smith, D.A. y Shukla, V.K. (2007). Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize: molecular and biochemical characterization. *Plant Physiol.* **144** (3): 1278-1291.
- Svergun D.I., Barberato, C. y Koch, M.H.J. (1995). CRY SOL - a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates *J. Appl. Cryst.* **28** 768-773.
- Sweetman D., Johnson, S., Caddick, S.E., Hanke, D.E. y Brearley, C.A. (2006). Characterization of an *Arabidopsis* inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). *Biochem J.* **394** (Pt 1): 95-103.
- Takazawa K. y Erneux, C. (1991). Identification of residues essential for catalysis and binding of calmodulin in rat brain inositol 1,4,5-trisphosphate 3-kinase. *Biochem J.* **280** ( Pt 1) 125-129.
- Takazawa K., Lemos, M., Delvaux, A., Lejeune, C., Dumont, J.E. y Erneux, C. (1990). Rat brain inositol 1,4,5-trisphosphate 3-kinase. Ca<sup>2+</sup> sensitivity, purification and antibody production. *Biochem J.* **268** (1): 213-217.
- Takazawa K., Perret, J., Dumont, J.E. y Erneux, C. (1991). Molecular cloning and expression of a human brain inositol 1,4,5-trisphosphate 3-kinase. *Biochem Biophys Res Commun.* **174** (2): 529-535.

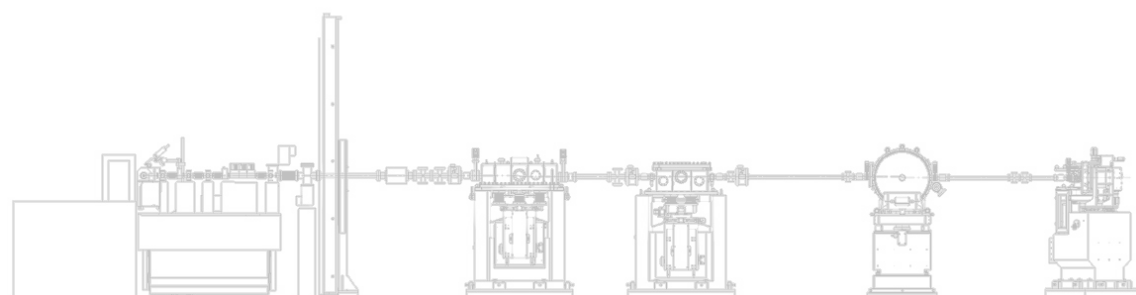
- Takazawa K., Perret, J., Dumont, J.E. y Erneux, C.** (1991). Molecular cloning and expression of a new putative inositol 1,4,5-trisphosphate 3-kinase isoenzyme. *Biochem J.* **278** ( Pt 3) 883-886.
- Tan X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M. y Zheng, N.** (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature.* **446** (7136): 640-645.
- Taylor D.A., Sack, J.S., Maune, J.F., Beckingham, K. y Quijcho, F.A.** (1991). Structure of a recombinant calmodulin from *Drosophila melanogaster* refined at 2.2-Å resolution. *J Biol Chem.* **266** (32): 21375-21380.
- Taylor G.** (2003). The phase problem. *Acta Crystallogr D Biol Crystallogr.* **59** (Pt 11): 1881-1890.
- Taylor S.S., Radzio-Andzelm, E., Madhusudan, Cheng, X., Ten Eyck, L. y Narayana, N.** (1999). Catalytic subunit of cyclic AMP-dependent protein kinase: structure and dynamics of the active site cleft. *Pharmacol Ther.* **82** (2-3): 133-141.
- Ten Eyck L. y Watenpaugh, K.D.** (2006). Introduction to refinement. *International Tables of Crystallography. Volumen F: Crystallography of biological macromolecules.* **18.** 369-374.
- Toker A.** (2002). Phosphoinositides and signal transduction. *Cell Mol Life Sci.* **59** (5): 761-779.
- Tronrud D.E.** (2004). Introduction to macromolecular refinement. *Acta Crystallogr D Biol Crystallogr.* **60** (Pt 12 Pt 1): 2156-2168.
- Tronrud D.E.** (2007). Introduction to macromolecular refinement. *Methods Mol Biol.* **364** 231-254.
- Vajanaphanich M., Schultz, C., Rudolf, M.T., Wasserman, M., Enyedi, P., Craxton, A., Shears, S.B., Tsien, R.Y., Barrett, K.E. y Traynor-Kaplan, A.** (1994). Long-term uncoupling of chloride secretion from intracellular calcium levels by Ins(3,4,5,6)P<sub>4</sub>. *Nature.* **371** (6499): 711-714.
- Vanweyenbergh V., Communi, D., D'Santos, C.S. y Erneux, C.** (1995). Tissue- and cell-specific expression of Ins(1,4,5)P<sub>3</sub> 3-kinase isoenzymes. *Biochem J.* **306** 429-435.
- Verbsky J., Lavine, K. y Majerus, P.W.** (2005). Disruption of the mouse inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression. *Proc Natl Acad Sci U S A.* **102** (24): 8448-8453.
- Verbsky J.W., Wilson, M.P., Kisseleva, M.V., Majerus, P.W. y Wente, S.R.** (2002). The synthesis of inositol hexakisphosphate. Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *J Biol Chem.* **277** (35): 31857-31862.
- Vetter S.W. y Leclerc, E.** (2003). Novel aspects of calmodulin target recognition and activation. *Eur J Biochem.* **270** (3): 404-414.
- Voglmaier S.M., Bembenek, M.E., Kaplin, A.I., Dorman, G., Olszewski, J.D., Prestwich, G.D. y Snyder, S.H.** (1996). Purified inositol hexakisphosphate kinase is an ATP synthase: diphosphoinositol pentakisphosphate as a high-energy phosphate donor. *Proc Natl Acad Sci U S A.* **93** (9): 4305-4310.
- Volkov V.V. y Svergun, D.I.** (2003). Uniqueness of ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* **36** 860-864.
- Vonrhein C., Blanc, E., Roversi, P. y Bricogne, G.** (2007). Automated structure solution with autoSHARP. *Methods Mol Biol.* **364** 215-230.



- Wallace A.C., Laskowski, R.A. y Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8** (2): 127-134.
- Waterhouse A.M., Procter, J.B., Martin, D.M., Clamp, M. y Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* **25** (9): 1189-1191.
- Weiss M.S. (2000). Global indicators of X-ray data quality. *J Appl Crystallogr.* **34** 130-135.
- Wilson M.P. y Majerus, P.W. (1996). Isolation of inositol 1,3,4-trisphosphate 5/6-kinase, cDNA cloning and expression of the recombinant enzyme. *J Biol Chem.* **271** (20): 11904-11910.
- Winn M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A. y Wilson, K.S. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr.* **67** (Pt 4): 235-242.
- Wood W.B. (1966). Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *J Mol Biol.* **16** (1): 118-133.
- Woodring P.J. y Garrison, J.C. (1997). Expression, purification, and regulation of two isoforms of the inositol 1,4,5-trisphosphate 3-kinase. *J Biol Chem.* **272** (48): 30447-30454.
- Xia H.J. y Yang, G. (2005). Inositol 1,4,5-trisphosphate 3-kinases: functions and regulations. *Cell Res.* **15** (2): 83-91.
- Xu R., Sen, N., Paul, B.D., Snowman, A.M., Rao, F., Vandiver, M.S., Xu, J. y Snyder, S.H. (2013). Inositol Polyphosphate Multikinase Is a Coactivator of p53-Mediated Transcription and Cell Death. *Sci Signal.* **6** (269): ra22.
- Yamniuk A.P. y Vogel, H.J. (2004). Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol Biotechnol.* **27** (1): 33-57.
- Yang X. y Shears, S.B. (2000). Multitasking in signal transduction by a promiscuous human Ins(3,4,5,6)P(4) 1-kinase/Ins(1,3,4)P(3) 5/6-kinase. *Biochem J.* **351** Pt 3 551-555.
- Yañez M., Gil-Longo, J. y Campos-Toimil, M. (2012). Calcium binding proteins. *Adv Exp Med Biol.* **740** 461-482.
- Yap K.L., Kim, J., Truong, K., Sherman, M., Yuan, T. y Ikura, M. (2000). Calmodulin target database. *J Struct Funct Genomics.* **1** (1): 8-14.
- York J.D. (2006). Regulation of nuclear processes by inositol polyphosphates. *Biochim Biophys Acta.* **1761** (5-6): 552-559.
- York J.D., Odom, A.R., Murphy, R., Ives, E.B. y Wente, S.R. (1999). A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science.* **285** (5424): 96-100.
- Zahn K. (1996). Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth. *J Bacteriol.* **178** (10): 2926-2933.
- Zeng L., Webster, S.V. y Newton, P.M. (2012). The biology of protein kinase C. *Adv Exp Med Biol.* **740** 639-661.
- Zhang K.Y.J., Cowtan, K., Main, P. y Read, R.J. (2006). Density modification and phase combination. *International Tables of Crystallography. Volumen F: Crystallography of Biological Macromolecules.* **15**. 311-331.

- Zhou G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W.R. y Chapman, M.S.** (1998). Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc Natl Acad Sci U S A.* **95** (15): 8449-8454.





## 7 ANEXOS



## 7.1 Anexo 1: resolución de estructuras de proteínas mediante cristalografía de rayos X

### 7.1.1 Obtención de muestras de proteínas

Para poder abordar el estudio cristalográfico de una proteína se requiere que la muestra cumpla ciertos requisitos, como son un alto grado de pureza (90-95%), homogeneidad, estabilidad y cantidad suficiente, del orden de miligramos (Bergfors, 2009).

El gran desarrollo que han tenido en los últimos 30 años las técnicas de expresión y purificación de proteínas recombinantes hace que estos requerimientos sean más sencillos de satisfacer. Los pasos básicos para la obtención de la muestra son la clonación, la expresión y la purificación (Hughes y Stock, 2006).

La clonación es el proceso por el cual el DNA codificante para la proteína de interés se introduce en un vector que permite la expresión de la proteína en diferentes organismos. La elección de vector va ligada a la elección del organismo en el cual se producirá la proteína y a la etiqueta que se quiere emplear para facilitar su posterior purificación. En algunas ocasiones se hacen necesarias modificaciones en proteínas ya clonadas. Una de las modificaciones más habituales es la introducción de mutaciones puntuales en la secuencia de aminoácidos (mutagénesis dirigida) (Papworth *et al.*, 1996). Una vez obtenido el vector de expresión, éste se introduce en el organismo seleccionado para la expresión de la proteína. Los organismos empleados varían, pudiendo ser desde bacterias (*Escherichia coli*, *Lactococcus* sp.) u hongos (*Saccharomyces cerevisiae*, *Pichia pastoris*) a sistemas más complejos, como células de insecto (por ejemplo células Sf9) o células de mamífero (por ejemplo células CHO).

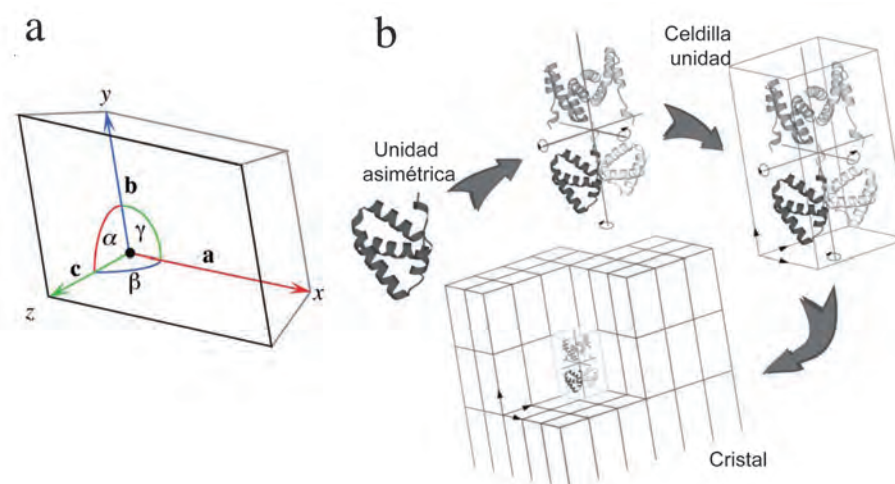
El proceso de purificación comienza con la extracción del contenido del organismo. Generalmente este paso se realiza mediante la ruptura (lisis) celular del cultivo, bien de forma mecánica mediante aplicación de presión (prensa de French), o bien con el empleo de ultrasonidos (sonicador). Posteriormente se realiza la clarificación, consistente en la eliminación de los restos celulares mediante centrifugación a altas velocidades. Finalmente se procede a la captura de la proteína, generalmente mediante cromatografías de afinidad. Para alcanzar un buen grado de pureza, se suelen combinar varios pasos cromatográficos, que en muchos casos, concluyen con una cromatografía de exclusión molecular.

## 7.1.2 Cristalización de proteínas

### 7.1.2.1 Introducción

Uno de los mayores obstáculos a la hora de determinar la estructura tridimensional de una proteína mediante cristalografía de rayos X es la obtención de cristales de proteína. Estos son imprescindibles debido a que solo la materia ordenada es capaz de difractar los rayos X.

Los cristales son el resultado de una agregación controlada de las moléculas en la que éstas se disponen de manera ordenada, simétrica y periódica. Este estado de agregación puede describirse como una red formada por la repetición, por traslación, de una celdilla unidad (**Figura 7.1a**) a lo largo de las tres direcciones del espacio (**Figura 7.1b**). Esta celdilla unidad está definida por 3 ángulos ( $\alpha, \beta, \gamma$ ) y tres ejes ( $a, b, c$ ) (**Figura 7.1a**). El contenido molecular de la celdilla unidad se obtiene mediante la aplicación de las operaciones de simetría de grupo espacial al motivo básico, denominado unidad asimétrica (**Figura 7.1b**), que a su vez puede contener una o más moléculas. Según las operaciones de simetría y el tipo de celdilla, los cristales pueden clasificarse en diferentes grupos espaciales. Existen un total de 230 grupos espaciales posibles, los cuales representan 230 modos diferentes de repetir un objeto en el espacio. En el caso de las proteínas, y debido a su quiralidad, no es posible la presencia de operaciones de simetría que impliquen centros y planos de simetría, reduciéndose el número grupos espaciales posibles a 65 (McPherson, 2009).



**Figura 7.1. El cristal.** a) Ejes y ángulos de una celdilla unidad. Figura tomada de Rhodes, (2006). b) Los cristales se forman a partir de una unidad mínima de repetición, denominada unidad asimétrica. La aplicación de los elementos de simetría de un determinado grupo espacial sobre ésta, generan la celdilla unidad. La traslación de la celdilla unidad en las tres direcciones del espacio da lugar al cristal. Figura modificada de McPherson, (2009).



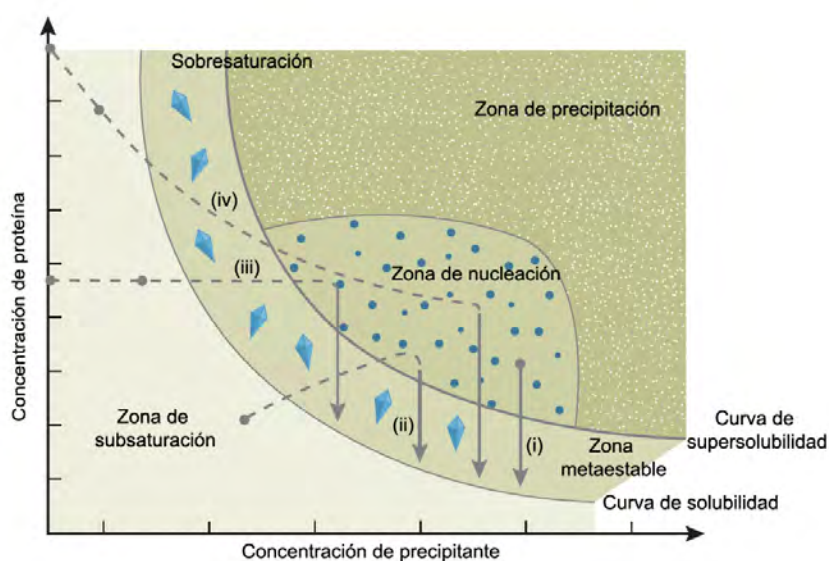
Los cristales de proteína suelen ser de un tamaño reducido, presentando celdillas unidad de gran tamaño (en algunos casos, varios cientos de angstroms). La combinación de estas dos circunstancias hace que los cristales de proteína difracten de forma mucho más débil los rayos X. Estos también presentan un elevado contenido de solvente (Matthews, 1968) y escasas interacciones entre las moléculas que forman el cristal, lo que hace que sean frágiles y muy sensibles a variaciones térmicas (McPherson, 2009). Por otro lado, la existencia de canales de solvente posibilita la introducción de moléculas de interés en el cristal.

#### 7.1.2.2 Base teórica del proceso de cristalización

El proceso de cristalización es un fenómeno de transición de fases y, como tal, puede ser explicado empleando un diagrama de fases (**Figura 7.2**). En el caso de la cristalización de proteínas, podemos dividir el diagrama en cuatro áreas (McPherson, 1999; Chayen y Saridakis, 2008):

- Un área de alta sobresaturación, donde la proteína precipita de forma rápida y no controlada, formando agregados amorfos (zona de precipitación).
- Un área de moderada sobresaturación, donde se produce el proceso de nucleación.
- Un área de baja sobresaturación, donde se produce el proceso de crecimiento cristalino (la denominada zona metaestable).
- Una zona de subsaturación, en la cual la proteína es totalmente soluble, y no es posible obtener cristales.

El proceso de cristalización consta, por tanto, de dos etapas: una primera de nucleación, y una segunda de crecimiento, que se produce una vez que el núcleo del cristal ha alcanzado un tamaño crítico.



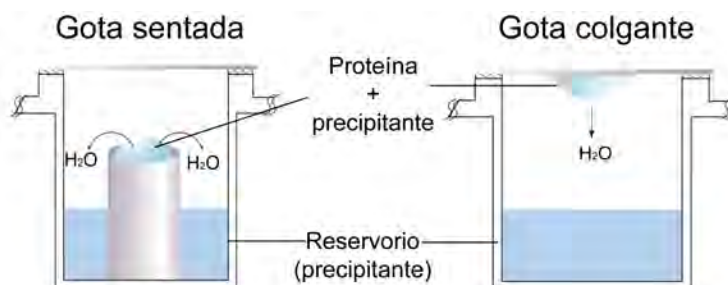
**Figura 7.2. Diagrama de fases del proceso de cristalización.** Las líneas representan de qué forma alcanza las zonas de nucleación y metaestable cada uno de los métodos de cristalización: i) *Microbatch*. ii) Difusión de vapor. iii) Diálisis. iv) Contradifusión a través de interfase. Imagen modificada de (Chayen y Saridakis, 2008).

### 7.1.2.3 Metodología de cristalización. Búsqueda y mejora de condiciones de cristalización

La obtención de una solución sobresaturada de proteína se realiza mezclando la solución de proteína concentrada con una solución precipitante. Existen tres grandes grupos de precipitantes: sales, polímeros orgánicos (como por ejemplo el polietilenglicol, PEG) y solventes orgánicos (alcoholes y polialcoholes) (Giegé y McPherson, 2006). Además de esto, otro factor importante a la hora de la cristalización de proteínas es el pH, el cual se regula usando diferentes soluciones tampón.

Se han desarrollado múltiples métodos para llevar a cabo el proceso de cristalización, como los métodos de difusión de vapor, el método de *microbatch*, la diálisis y la contradifusión a través de interfase. Cada uno de estos métodos implica que se alcance la zona de nucleación y la zona metaestable de forma diferente (**Figura 7.2**) (McPherson, 1999; Chayen y Saridakis, 2008). De todos estos métodos, el más utilizado es el de difusión de vapor. En esta técnica, una pequeña gota compuesta por una mezcla de proteína y precipitante se equilibra contra un reservorio que contiene solamente la solución precipitante. De este modo se produce un flujo neto de vapor de agua en la dirección del reservorio hasta que se alcance el equilibrio. Esta fuga de agua hace que se alcancen condiciones de sobresaturación en la gota (por el aumento de concentración del precipitante en la gota), lo que lleva a la proteína a precipitar o si las condiciones del experimento son las adecuadas, a la formación de núcleos y posteriormente de cristales.

Este método presenta dos variantes (**Figura 7.3**): la gota sentada (en la cual la gota descansa sobre un soporte) y la gota colgante (en la cual la gota está suspendida directamente sobre el precipitante en un cubreobjetos) (McPherson, 1999; Bergfors, 2009).



**Figura 7.3.** Variantes del método de cristalización por difusión de vapor.

La búsqueda de condiciones de cristalización para una proteína es un proceso empírico. El proceso de cristalización de cualquier proteína se divide en dos fases principales, una primera fase de búsqueda inicial de las condiciones de cristalización y una segunda de optimización (McPherson, 2004).

La búsqueda inicial de condiciones de cristalización se realiza actualmente mediante “técnicas de alto rendimiento” empleando robots especializados para la cristalización de proteínas (**Figura 7.4**) (Stevens, 2000). Este enfoque presenta múltiples ventajas: permite el empleo de menor cantidad de proteína, pudiéndose explorar un número de condiciones mucho mayor. En paralelo, se dispone también de una batería de soluciones comerciales que intentan abarcar el mayor número de condiciones posibles. Dichas soluciones están basadas en diferentes estrategias, aunque la más empleada es la de la matriz de dispersión (Jancarik y Kim, 1991). La idea básica es el cribado del espacio de cristalización mediante la combinación de condiciones con las que se han obtenido previamente cristales de otras proteínas.



**Figura 7.4.** Cristalización de alto rendimiento. Robots de cristalización del Departamento de Cristalografía y Biología Estructural del Instituto de Química Física Rocasolano.

Con el objetivo de mejorar la calidad de los cristales, las condiciones de cristalización obtenidas son optimizadas, escalándose a volúmenes mayores los experimentos de cristalización. Existe una gran variedad de técnicas de optimización: ajuste del pH, de la concentración del precipitante, de la concentración de proteína y de la temperatura. Otros recursos empleados son el uso de compuestos suplementarios, (moléculas de pequeño tamaño o detergentes) (McPherson, 1999; Bergfors, 2009; Rupp, 2009), así como la inclusión de los ligandos (sustratos, productos o cofactores fisiológicos de la proteína). También puede emplearse la técnica del sembrado, que consiste en la introducción de núcleos cristalinos, de forma que el cristal puede crecer a partir de estos núcleos, facilitándose así el proceso de crecimiento cristalino (Stura y Wilson, 1992; Bergfors, 2003). De entre todos los tipos de sembrado, el más empleado es el denominado *streak seeding*, consistente en la introducción de los núcleos cristalinos en la gota mediante una fibra natural.

En ocasiones se hace necesario obtener complejos de la proteína de interés unida a cierto ligando. Existen dos métodos para la obtención de complejos proteína-ligando: la cocrystalización, que consiste en la cristalización de la proteína en presencia del ligando (bien por mezcla en la gota o por incubación previa) y la inmersión, que se basa en inmersión de los cristales de proteína previamente obtenidos en soluciones que contienen el ligando.

### 7.1.3 Recogida de datos de difracción

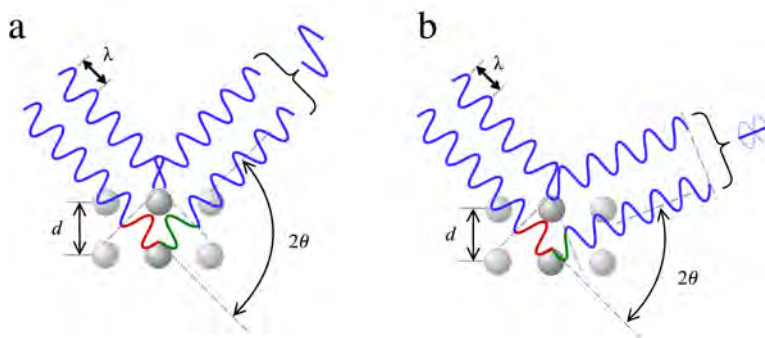
#### 7.1.3.1 El fenómeno de la difracción

La radiación electromagnética puede interactuar con la materia, dando lugar a multitud de fenómenos como la reflexión, la refracción, la dispersión, la absorción o la polarización de la luz. Uno de estos fenómenos es la difracción, fenómeno a través del cual se manifiesta la interacción de los rayos X con los cristales (u otra materia ordenada). Cuando un haz de rayos X incide sobre un cristal, los fotones interactúan con los electrones de los átomos del cristal, haciéndolos vibrar y convirtiéndolos a su vez en emisores secundarios de rayos X (dispersión). Cada electrón dispersa la radiación X en todas las direcciones del espacio. Una de las formas más intuitivas para explicar el fenómeno es la denominada Ley de Bragg (Bragg y Bragg, 1913; Giacovazzo, 2011) (**Figura 7.5**). Según ésta, los rayos X incidentes son reflejados por una serie de planos paralelos imaginarios que intersectan los diferentes átomos del cristal. Cada familia de planos, separados por una distancia  $d$ , se identifican por tres números enteros ( $hkl$ ) denominados índices de Miller. Cuando el ángulo de incidencia del haz ( $\theta$ ) entre uno de estos planos y el siguiente hace que la diferencia de camino recorrido entre los haces reflejados sea un múltiplo entero de la longitud de onda del rayo incidente, se produce interferencia constructiva con lo que se observa el fenómeno

de la difracción. Si esto no se cumple, las interferencias serán destructivas (**Ecuación 7.1**).

$$n\lambda = 2d \sin \theta$$

**Ecuación 7.1. Ley de Bragg.**



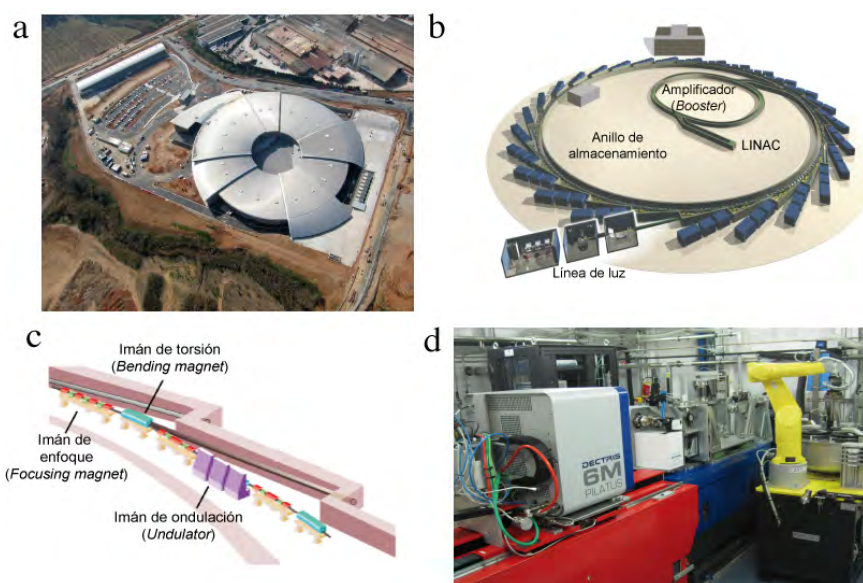
**Figura 7.5. Ley de Bragg.** a) Esquema que muestra que cuando la diferencia de camino recorrido entre los haces difractados por una familia de planos es un múltiplo entero de  $\lambda$  (1 en este caso) se producen interferencias constructivas y por tanto se observa difracción. b) Cuando el múltiplo de  $\lambda$  no es entero, se producen interferencias destructivas.

### 7.1.3.2 Generación de rayos X

Los rayos X son radiaciones electromagnéticas con una longitud de onda de entre 1000 y 0.1 Å, aunque normalmente la cristalografía emplea rayos X de una longitud de onda de aproximadamente 1-2 Å. Actualmente, se pueden emplear los rayos X producidos en equipos de ánodo rotante, o emplear la radiación producida en una estación sincrotrón.

La radiación sincrotrónica se genera en aceleradores de partículas en forma de anillo (**Figura 7.6a** y b), en los que un haz de partículas cargadas se encuentra circulando a velocidades cercanas a las de la luz en un campo magnético. Cuando este haz de electrones sufre desviaciones en su trayectoria debido a la presencia de imanes (**Figura 7.6c**), se emite una radiación electromagnética tangencial a la dirección de las partículas aceleradas. Esta radiación comprende un amplio espectro de longitudes de onda (la denominada radiación sincrotrónica), entre ellas rayos X de alta energía los cuales son empleados para la realización de los experimentos en las denominadas “líneas de luz” (**Figura 7.6d**). El uso de la radiación sincrotrónica tiene dos principales ventajas: un mayor brillo (la intensidad y enfoque del haz es mayor), lo que permite la obtención de datos de mayor calidad, el uso de cristales más pequeños, y la posibilidad de variar la longitud de onda, lo que permite emplear métodos basados en el fenómeno de la dispersión anómala para resolver el problema de la fase (Helliwell, 2006).





**Figura 7.6. Generación y uso de radiación de sincrotrón.** a) Vista aérea del sincrotrón ALBA (Barcelona), en la cual se aprecia la forma de anillo de un sincrotrón. Imagen tomada de <http://www.pemb.cat/en/?projecte=alba-synchrotron-light-source>. b) Esquema general de las diferentes partes de un sincrotrón. Imagen tomada y modificada de (Sorensen *et al.*, 2006). c) Tipos de imanes presentes en el anillo de almacenamiento. Imágenes tomadas de <http://www.esrf.eu/about/synchrotron-science/synchrotron>. d) Aspecto de la cabina experimental de una línea de luz dedicada a proteínas (cabina experimental de XALOC, sincrotrón ALBA).

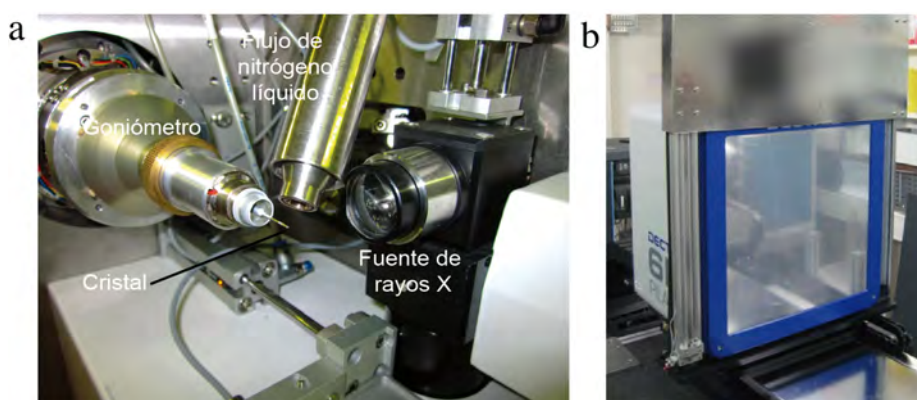
### 7.1.3.3 Recogida de datos de difracción

Los cristales de proteína pueden sufrir daños durante la recogida de datos, por lo cual es habitual recoger los datos de difracción a baja temperatura (Dauter y Wilson, 2006). Para ello, el cristal debe ser protegido. Entre los crioprotectores más comunes se encuentran el etilenglicol, el glicerol, el PEG de diferentes pesos moleculares, la sacarosa o algunas sales a altas concentraciones. Para crioproteger los cristales, éstos se ponen en contacto durante unos segundos con la solución crioprotectora, que suele estar formada por condiciones similares a las de cristalización, e incluyen diferentes porcentajes del agente crioprotector. Una vez crioprotegido, el cristal es cuidadosamente colocado en un lazo de nylon y congelado, quedando listo para el proceso de recogida de datos (Garman y Schneider, 1997; Pflugrath, 2004).

En la recogida de datos de difracción se hace incidir un haz monocromático de rayos X de alta energía con una longitud de onda de aproximadamente  $1\text{\AA}$  sobre el cristal y se recogen los haces difractados en un detector situado perpendicularmente al haz de rayos X (**Figura 7.7a**). Con el fin de obtener un conjunto de datos completo el cristal se rota sobre un eje. Este método, el más empleado actualmente, se denomina método de la rotación.

Por tanto, los elementos básicos que se requieren para realizar el experimento son: (Dauter y Wilson, 2006) (**Figura 7.7a**):

- **Fuente de rayos X:** el requerimiento principal es que sea de alta energía. (**Figura 7.7a**).
- **Detector:** permite registrar las intensidades y digitalizarlas para procesarlas posteriormente. Los más usados los de tipo Image Plate y los basados en las tecnologías CCD y CMOS (estos últimos denominados *pixel detectors*) (Gruner *et al.*, 2006; Kraft *et al.*, 2009) (**Figura 7.7b**)
- **Sistema de rotación del cristal:** consiste en una cabeza goniométrica con un motor asociado, que permite montar el cristal y realizar sucesivos giros del mismo (**Figura 7.7a**).
- **Sistema criogénico:** flujo de nitrógeno frío que mantiene la muestra a baja temperatura durante el transcurso de la medida (**Figura 7.7a**).



**Figura 7.7. Elementos para la realización de un experimento de difracción.** a) Entorno de la muestra. Imagen tomada en la línea XALOC del sincrotrón ALBA. b) Detector Pilatus 6M (Dectris Ltd), instalado en la línea ID29 del ESRF, Grenoble (Francia).

### 7.1.4 Procesado de datos de difracción

El resultado del experimento de difracción es un conjunto de imágenes digitales en el que se encuentran registradas las intensidades difractadas por el cristal (**Figura 7.8a**). Podemos dividir este procesado en varias fases: indexado e integración, escalado y finalmente, cálculo del módulo de los factores de estructura.

- **Indexado e integración:** durante éste este paso del procesado se analiza la geometría del espectro de difracción, lo que permite la determinación de los parámetros del cristal (orientación, celdilla unidad y grupo espacial más probable). El indexado propiamente dicho consiste en asignar índices  $hkl$  a las distintas reflexiones recogidas (producidas por un conjunto de planos de átomos  $hkl$ ). En el proceso de integración se calcula un valor para la intensidad  $I(hkl)$  de

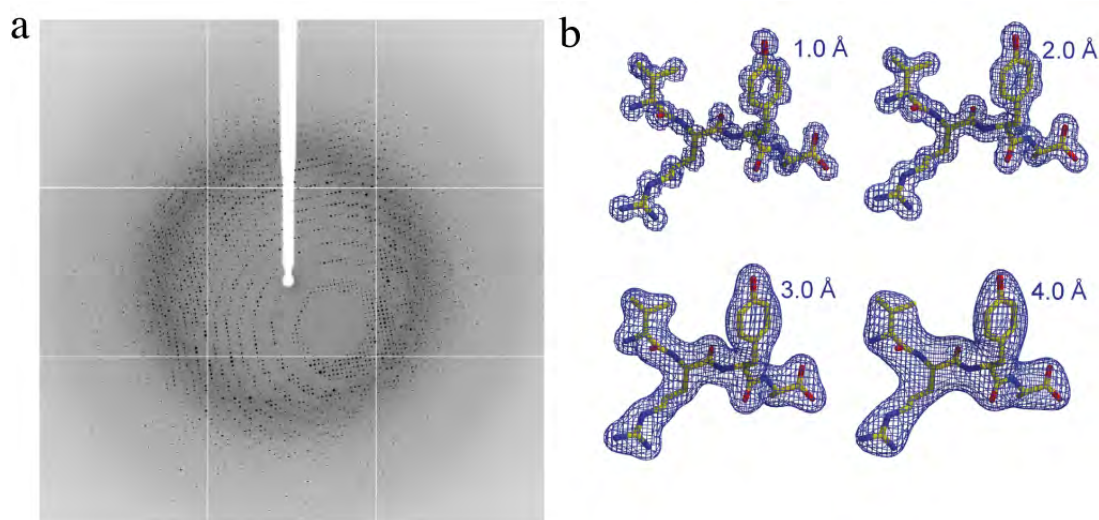


cada una de las reflexiones medidas, junto con una desviación estándar asociada  $\sigma I(hkl)$ . Durante este proceso también se realiza una estimación de las imperfecciones del cristal (mosaicidad) (Leslie, 2006; Rossmann, 2006).

- Escalado: este proceso consiste en la corrección de las intensidades obtenidas, teniendo en cuenta parámetros como el ruido de fondo de la imagen, la mosaicidad, o el decaimiento de la difracción durante el proceso de recogida. A su vez se realiza un promediado de las intensidades equivalentes (relacionadas por simetría) y se normalizan calculándose una desviación estándar (Evans, 2006; Kabsch, 2006). Del proceso de escalado se obtienen unos parámetros que son empleados para evaluar la calidad general de los datos obtenidos (Weiss, 2000; Evans, 2011). Algunos de estos parámetros son:
  - ❖  $R_{\text{merge}}/R_{\text{sym}}$ : factor de desacuerdo de las intensidades de reflexiones relacionadas por simetría. Un buen conjunto de datos debería presentar un  $R_{\text{merge}}$  menor del 10%-15%, y del 50%-60% en la última capa de resolución.
  - ❖  $I/\sigma(I)$  promedio: representa la relación de intensidad de la señal con respecto al ruido de las reflexiones promediadas. Valores mayores de 1.5-2 son aceptables (Evans, 2011).
  - ❖ Completitud: Porcentaje de reflexiones recogidas sobre el total posible. Es necesaria una completitud del orden de 90%-100% y al menos 80% en la capa de mayor resolución.
  - ❖ Multiplicidad: número de observaciones para cada reflexión. Cuanto mayor es este parámetro mejor es la estimación de los errores experimentales y sistemáticos. Es recomendable realizar la medida de cada reflexión como mínimo 4 veces.

Considerando el balance entre todos estos parámetros puede determinarse la resolución máxima de los datos de difracción. Cuanto mejor es la resolución, mayor definición tendrán los mapas de densidad electrónica, permitiendo una mejor determinación de la estructura (**Figura 7.8b**).

- Cálculo del módulo del factor de estructura o  $|F(hkl)|$ : el cual es proporcional a la raíz cuadrada de la intensidad de la difracción (ver Sección 3.1.6)



**Figura 7.8. Patrón de difracción. Relación entre la resolución y la calidad de los mapas de densidad electrónica.** a) Patrón de difracción de una proteína, recogido empleando radiación de sincrotrón. b) Influencia de la resolución en el nivel de detalle con que se puede construir el modelo estructural. Imagen de (Rupp, 2009).

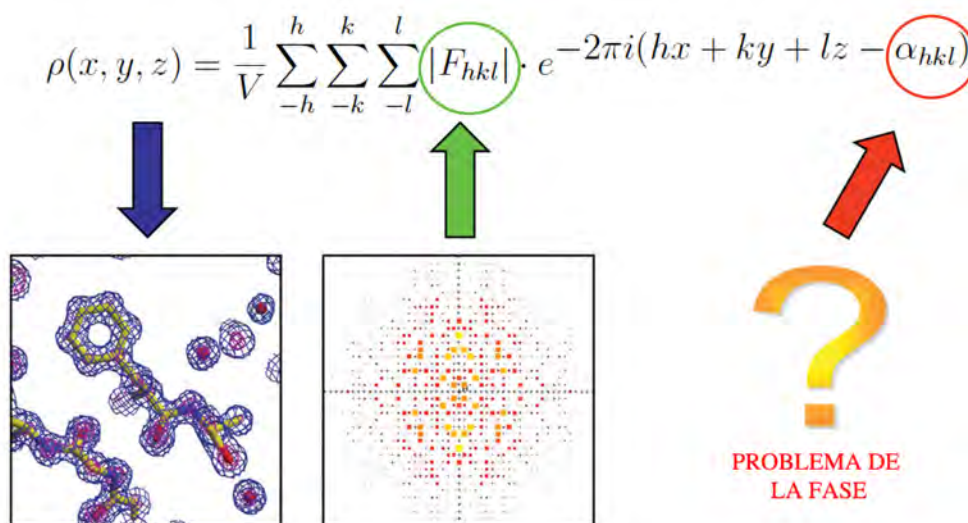
## 7.1.5 Resolución estructural

### 7.1.5.1 Introducción: el problema de la fase

La relación entre la estructura cristalina (en realidad su densidad electrónica,  $\rho(xyz)$ ), y el experimento de difracción, o lo que es lo mismo, entre espacio real y recíproco, se establece mediante la transformada de Fourier. Así pues (**Ecuación 7.2**) la densidad electrónica,  $\rho(xyz)$ , en cada punto  $(xyz)$  del cristal está relacionada con  $F_{hkl}$  (el factor de estructura). Los factores de estructura son magnitudes vectoriales que por tanto vienen caracterizados por un módulo  $|F_{hkl}|$ , obtenido experimentalmente a partir de las intensidades observadas en el patrón de difracción y una fase  $\alpha_{hkl}$ , cuya información se pierde durante el experimento de difracción, dando lugar al denominado problema de la fase (**Figura 7.9**) (Taylor, 2003; Rupp, 2009). Como se verá, la información sobre las fases puede recuperarse empleando diferentes aproximaciones experimentales.

$$\rho(x, y, z) = \frac{1}{V} \sum_{-h}^h \sum_{-k}^k \sum_{-l}^l |F_{hkl}| \cdot e^{-2\pi i(hx + ky + lz - \alpha_{hkl})}$$

**Ecuación 7.2. Función de densidad electrónica.**



**Figura 7.9. El problema de la fase.** El cálculo de la densidad electrónica de una proteína requiere dos tipos de datos: amplitudes y fases de los vectores asociados a las ondas difractadas. Las amplitudes pueden obtenerse empleando las intensidades registradas en el patrón de difracción mientras que la información relativa a las fases se pierde durante el experimento. Esta pérdida de información provoca lo que se ha denominado “el problema de la fase”. Imagen modificada de Rupp (2009).

#### 7.1.5.2 Métodos experimentales de obtención de las fases

Existen tres aproximaciones para obtener dicha información:

- **Método del reemplazamiento isomorfo simple/múltiple (SIR/MIR):** los métodos de reemplazamiento isomorfo se basan en introducir en el cristal átomos con alto poder dispersor de los rayos X (es decir átomos con números atómicos elevados), sin que ello distorsione la estructura cristalina, con el objetivo de localizar su posición y emplearla para hacer una estimación inicial de las fases (mediante la comparación entre conjuntos de datos que contengan el átomo pesado y otros que no contengan dicho átomo). Fue el primer método desarrollado para la resolución de estructuras de proteínas (Green *et al.*, 1954; Carvin *et al.*, 2006).
- **Método de la dispersión anómala simple/múltiple (SAD/MAD):** se basan en la introducción en el cristal de átomos que sean dispersores anómalos de los rayos X. Para localizar su posición se usa la diferencia en la dispersión que producen estos átomos a la misma (SAD) o diferentes (MAD) longitudes de onda. Estos métodos fueron aplicados por primera vez con éxito durante los años ochenta (Hendrickson y Teeter, 1981; Hendrickson *et al.*, 1985; Kahn *et al.*, 1985).
- **Método de Reemplazamiento Molecular (MR):** se basa en calcular las fases de la proteína de interés empleando para ello las fases de una proteína homóloga

como modelo. Fue desarrollado por Michael Rossmann y David Blow durante los años sesenta (Rossmann y Blow, 1962).

Los métodos empleados para la resolución de las estructuras de este trabajo fueron el método MAD y el Reemplazamiento Molecular.

#### 7.1.5.2.1 Método de dispersión anómala múltiple (MAD)

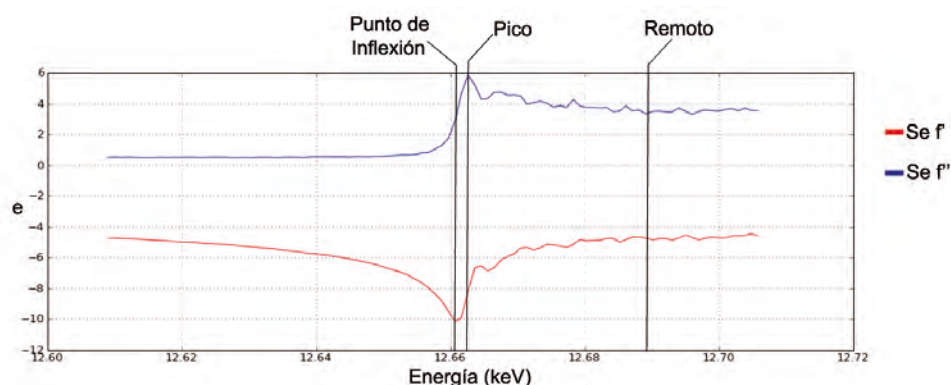
Como se dijo previamente, la interacción de los rayos X con la materia cristalina puede producir fenómenos diferentes de la difracción, como por ejemplo absorción y dispersión. La dispersión atómica de un átomo viene caracterizada por el denominado factor de dispersión atómico, que consta de un componente “normal” y un componente “anómalo”, éste último compuesto de dos partes, una real ( $f'$ ), que da cuenta de los efectos de dispersión y una imaginaria ( $if''$ ), que refleja los efectos de absorción (**Ecuación 7.3**). Los átomos pueden absorber radiación a determinadas longitudes de onda, dispersándola a su vez. Sin embargo, esta dispersión sufre un desplazamiento en fase y amplitud, fenómeno conocido como dispersión anómala. Átomos con un elevado número de electrones (átomos pesados), presentan fuertes efectos de dispersión anómala, mientras que en los más ligeros (H, C, N, O) la dispersión anómala es relativamente escasa. La inclusión de estos dispersores anómalos en el cristal se suele realizar mediante cocrystalización o inmersión en soluciones que los contengan. Actualmente, también existen técnicas de expresión de proteínas que permiten sustituir el azufre de las cisteínas y metioninas por selenio (Hendrickson *et al.*, 1990).

$$f = f_0 + f' + if''$$

**Ecuación 7.3. Factor de dispersión atómico.**

La utilización de este fenómeno es más eficaz cuando comparamos la difracción de un cristal a distintas longitudes de onda. La dependencia del fenómeno de la longitud de onda causa que este tipo de experimentos sólo puedan llevarse a cabo en lugares donde se tiene un control preciso de ésta, como son las instalaciones de radiación sincrotrónica. Desde el punto de vista práctico, el experimento se realiza tomando varios (3 o 4) conjuntos de datos a diferentes longitudes de onda. Las longitudes de onda se escogen con el objetivo de maximizar los efectos de absorción y dispersión. Así pues se recoge un primer conjunto de datos a una longitud de onda ( $\lambda_1$ ) coincidente con el pico de absorción del átomo pesado que contenga el cristal (máximo valor de  $f''$ ). A continuación se recoge un segundo conjunto a una longitud de onda ( $\lambda_2$ ) en el punto de inflexión de la curva de absorción  $f''$ , el cual coincide con la el punto en el que la dispersión es mínima (mínimo de la curva  $f'$ ). Para finalizar se recoge un tercer (y en ocasiones también un cuarto) conjunto de datos a una longitud de onda alejada de las dos anteriores (remoto,  $\lambda_3/\lambda_4$ ) (**Figura 7.10**). Las longitudes de onda exactas pueden ser

determinadas mediante un barrido de la fluorescencia del cristal en torno al pico de absorción del átomo pesado en cuestión, ya que el entorno de éstos puede variar las longitudes de onda reales para la dispersión anómala de cada átomo.



**Figura 7.10. Espectro de fluorescencia alrededor del pico de absorción del selenio.** Gracias al espectro pueden escogerse las longitudes de onda adecuadas para la realización de un experimento de MAD.

El efecto de la dispersión anómala es también observable al comparar las intensidades de parejas de reflexiones simétricas (pares de Friedel). Estos pares deberían ser iguales en intensidad pero con fases de diferente signo. No obstante, la inclusión de átomos pesados rompe esta igualdad ( $I(hkl) \neq I(-h-k-l)$ ), dando lugar a las diferencias anómalas o de Bijvoet.

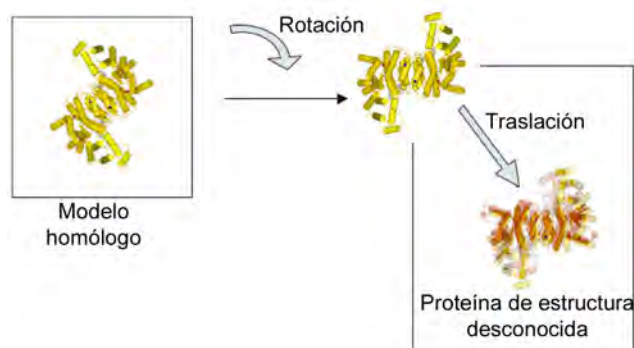
El análisis de la señal anómala obtenida permite localizar la posición de los átomos pesados, utilizando métodos de Patterson o bien métodos directos (Rupp, 2009; Giacovazzo, 2011). A partir de estas posiciones, se puede estimar el conjunto de fases iniciales que permite el cálculo de un mapa de densidad.

#### 7.1.5.2.2 Método de Reemplazamiento Molecular (MR)

Cuando se dispone de la estructura de una proteína con secuencia homóloga a la que se quiere resolver, ésta puede emplearse como modelo para el cálculo de unas fases iniciales. Para ello se requiere que el modelo tenga una identidad de secuencia con nuestra proteína de al menos un 25-30% (Rupp, 2009).

El método se basa en el cálculo de una función de rotación [R] (orientación) y una función de traslación [T] (posición en la celdilla) para el modelo, de forma que se “sitúa” el modelo en el lugar de la celdilla donde se encuentra nuestra proteína (**Figura 7.11**). Una vez encontrada la mejor rotación y traslación se calculan las fases del modelo posicionado en nuestra celdilla y se combinan con las amplitudes

experimentales para calcular un mapa inicial de densidad electrónica que permita trazar la cadena de polipeptídica.



**Figura 7.11. Método del Reemplazamiento Molecular.** Figura modificada de (Taylor, 2003).

### 7.1.5.3 Mejora de fases

En muchas ocasiones, las fases obtenidas no son lo suficientemente precisas como para permitir la construcción del modelo. En estos casos se emplean métodos de modificación de densidad para mejorar la calidad de los mapas obtenidos. Existen diferentes técnicas para esto: el aplanado de solvente (*solvent flattening*), el promediado por simetría no cristalográfica (*non-crystallographic averaging*) y el ajuste a histograma (*histogram matching*). El proceso de mejora de fases es cíclico (Zhang *et al.*, 2006; Cowtan, 2010).

## 7.1.6 Construcción, refinado y validación del modelo

### 7.1.6.1 Construcción y refinado del modelo

Las fases obtenidas durante el proceso de resolución estructural permiten calcular un mapa de densidad electrónica, a partir del cual se construye un modelo inicial de la proteína. Para ello se emplean programas gráficos de modelado que de forma interactiva permiten ir introduciendo cambios en el modelo, ampliándolo y ajustándolo al mapa de densidad electrónica. Una vez que se ha construido el modelo, se realiza un refinamiento matemático del mismo en el que se ajustan diferentes parámetros con el objetivo de optimizar el acuerdo entre los factores de estructura experimentales y los calculados a partir del modelo. El número de parámetros experimentales obtenido suele ser menor que el número de parámetros a refinar (Tronrud, 2004; Ten Eyck y Watenpugh, 2006). Para solucionar este problema, se introducen una serie de restricciones geométricas sobre el modelo, basándose en datos experimentales conocidos como las distancias entre átomos, ángulos de enlace o torsiones (Engh y Huber, 1991). Los parámetros que se refinan durante el proceso son las coordenadas



atómicas ( $x,y,z$ ), las cuales indican la posición de cada átomo dentro de la celdilla unidad, los factores térmicos ( $B$ ), los cuales dan una idea de la agitación térmica o movilidad de cada átomo, y la ocupación o fracción de celdillas unidad en las que un átomo ocupa una determinada posición.

Existen distintos métodos de refinamiento. En las primeras etapas se realiza refinamiento de cuerpo rígido (*Rigid body*) y calentamiento simulado (*Simulated annealing*), mientras que en etapas posteriores, se utilizan los métodos probabilísticos de máxima verosimilitud (*Maximun likelihood*) (Tronrud, 2004; Tronrud, 2007). En los últimos estadios del refinamiento, puede emplearse el método de TLS (*Translation/Libration/Screw*). Este método se basa en considerar ciertos grupos de átomos como un cuerpo rígido cuyo movimiento se describe mediante el uso de tres matrices (Painter y Merritt, 2006).

El proceso de refinado y reconstrucción debe asegurar que los cambios introducidos conduzcan hacia un modelo lo más parecido posible a la estructura real de la molécula. Para ello se introdujo un parámetro estadístico, el factor de desacuerdo  $R_{\text{work}}$  (**Ecuación 7.4**) que estima la calidad general del modelo. Este indicador nos da las diferencias entre los factores de estructura observados ( $|F^{\text{obs}}|$ ) y los calculados a partir el modelo ( $|F^{\text{calc}}|$ ). Este valor debe disminuir a lo largo del proceso de refinamiento. El valor es dependiente de la resolución; por ejemplo, con un conjunto de datos a 2.5 Å el valor  $R_{\text{work}}$  debería situarse en torno al 25%.

$$R_{\text{work}}/R_{\text{free}} = \frac{\sum_{hkl} |F_{hkl}^{\text{obs}} - F_{hkl}^{\text{calc}}|}{\sum_{hkl} F_{hkl}^{\text{obs}}}$$

**Ecuación 7.4.**  $R_{\text{work}}/R_{\text{free}}$ . El cálculo del  $R_{\text{free}}$  se basa en la exclusión de un 5-10% de las reflexiones observadas.

No obstante, puede ocurrir que el  $R_{\text{work}}$  disminuya sin que el modelo mejore, lo que se conoce como “sobrefinamiento”. Para evitar este problema se implementó un método de validación cruzada (Brunger, 1992). En este método se calcula un segundo parámetro, el  $R_{\text{free}}$  (para su cálculo se emplea la misma expresión que en la **Ecuación 7.4**). El  $R_{\text{free}}$  es un factor de desacuerdo similar al anterior, pero en este caso calculado con un 5-10% de reflexiones, que son excluidas del refinamiento. Este valor actúa por tanto como control contra el sobrefinamiento, pues sólo los cambios que disminuyan el  $R_{\text{work}}$  y el  $R_{\text{free}}$  serán correctos. El refinamiento se da por finalizado cuando el modelo se ajusta al mapa de densidad electrónica y los factores de desacuerdo convergen.



#### 7.1.6.2 Validación

La última fase del proceso de determinación estructural consiste en la validación del modelo obtenido. Para ello se comprueba que los diferentes parámetros del modelo cumplan con las distintas restricciones geométricas (ángulos y distancias de enlace entre átomos) y estereoquímicas (ángulos phi ( $\phi$ ) y psi ( $\psi$ ) de los diferentes residuos en la cadena polipeptídica). Estas últimas son evaluadas mediante el diagrama de Ramachandran (Ramachandran *et al.*, 1963). A su vez, se analiza que no existan distancias entre átomos por debajo de los parámetros establecidos (Kleywegt, 2000).

El resultado final del proceso, consistente en las coordenadas atómicas del modelo y los correspondientes factores de estructura, se deposita en el Banco de Datos de Proteínas (PDB, *Protein Data Bank*) (Berman *et al.*, 2003), en donde son evaluados y validados, para finalmente asignarles un código alfanumérico (código PDB) que identificará la estructura de la proteína.

## 7.2 Anexo 2: materiales empleados en la elaboración de este trabajo

### 7.2.1 Listado de Oligonucleótidos

Número	Secuencia	Propósito	Proteína
1	CGAGCTCATGGAGATGATT TTGGAGG	Introducción diana SacI	IP <sub>5</sub> 2-K
2	CCGCTCGAGTTAGCTGTGG GAAGGTTTGTAGTTGC	Introducción diana XhoI	IP <sub>5</sub> 2-K
3	GGAGATATACATATGGAGA TGATTTTGGAGGAG	Introducción diana NdeI	IP <sub>5</sub> 2-K
4	GGGGTACCTTAGCTGTGGG AAGGTTTGTAGTTGC	Introducción diana KpnI	IP <sub>5</sub> 2-K
5	GTGGGGAAAGTATTGTTA TACAGAAGGCTCGGAGG	Mutación R40V	IP <sub>5</sub> 2-K <sub>R40V</sub>
6	CCTCCGAGCCTTCTGTATAA CAATCACTTTCCCCAC	Mutación R40V	IP <sub>5</sub> 2-K <sub>R40V</sub>
7	CGTCCGCTATGGATTGTAA TGCAGC	Mutación R130I	IP <sub>5</sub> 2-K <sub>R130I</sub>
8	GCTGCATTAACAATCCATAG CGGACG	Mutación R130I	IP <sub>5</sub> 2-K <sub>R130I</sub>
9	CATTAGTGTGAAATAGCGC CTAAATGCGGATTTCTTCC	Mutación K168A	IP <sub>5</sub> 2-K <sub>K168A</sub>
10	GGAAGAAATCCGCATTTAG GCGCTATTTC AACACTAATG	Mutación K168A	IP <sub>5</sub> 2-K <sub>K168A</sub>
11	CATTAGTGTGAAATAAATC CTAAATGCGGATTTCTTCC	Mutación K168N	IP <sub>5</sub> 2-K <sub>K168N</sub>
12	GGAAGAAATCCGCATTTAG GATTTATTTC AACACTAATG	Mutación K168N	IP <sub>5</sub> 2-K <sub>K168N</sub>
13	GAAATAAAGCCTAGCTGCG GTTTCTTCC	Mutación D368A	IP <sub>5</sub> 2-K <sub>D368A</sub>
14	GGAAGAAATCCGCAGCTAG GCTTTATTTC	Mutación D368A	IP <sub>5</sub> 2-K <sub>D368A</sub>
15	CAAGGTACATTTTCATTGCTC TAAGCCTGAAACC	Mutación D407A	IP <sub>5</sub> 2-K <sub>D407A</sub>
15	GGTTTCAGGCTTAGAGCAAT GAAATGTACCTTG	Mutación D407A	IP <sub>5</sub> 2-K <sub>D407A</sub>
17	CCATTCTCTTAAGTGGTGCC AGGCTTAGATCAATG	Mutación K411A	IP <sub>5</sub> 2-K <sub>K411A</sub>
18	CATTGATCTAAGCCTGGCAC CACTTAAGAGAATGG	Mutación K411A	IP <sub>5</sub> 2-K <sub>K411A</sub>
19	CCACTCTCAAGCCAATTTTC CGCGTATTC	Mutación N238A	IP <sub>5</sub> 2-K <sub>N238A</sub>
20	GAATACGCGGAAATTGGCT TGAGGAGTGG	Mutación N238A	IP <sub>5</sub> 2-K <sub>N238A</sub>
21	CCAAACAAGGAAGTTCTTG CACAAAGATATGTTTCTG	Mutación E85A	IP <sub>5</sub> 2-K <sub>E85A</sub>
22	CTGAACATATCTTTGTGCAA GAACTTCCTGTTTGG	Mutación E85A	IP <sub>5</sub> 2-K <sub>E85A</sub>
23	CAACCAGCCTGCCCCTATAT GTAAA	Mutación C330A	IP <sub>5</sub> 2-K <sub>C330A</sub>
24	TTTACATATAGGGGAGGCT GGTTG	Mutación C330A	IP <sub>5</sub> 2-K <sub>C330A</sub>
25	GAAATAAAGCCTAGCTGCG GATTTCTTCC	Mutación K170S	IP <sub>5</sub> 2-K <sub>K170S</sub>
26	GGAAGAAATCCGCAGCTAG GCTTTATTTC	Mutación K170S	IP <sub>5</sub> 2-K <sub>K170S</sub>
27	CCTTGCCCTATATCTAAAGA AGGTAGG	Mutación C333A	IP <sub>5</sub> 2-K <sub>C333A</sub>
28	CCTACCTCTTTAGATATAG GGCAAGG	Mutación C333A	IP <sub>5</sub> 2-K <sub>C333A</sub>

29	GCGGAATTGTCTCTAAATGC TTTACCTTTAGATG	Mutación H346A	IP <sub>5</sub> 2-K <sub>H346Q</sub>
30	CATCTAAAGGTAAAGCATT AGAGACAATTCCGC	Mutación H346A	IP <sub>5</sub> 2-K <sub>H346Q</sub>
31	CGTCCGCTAGCGCGTGTTAA TGC	Mutación W129A	IP <sub>5</sub> 2-K <sub>W129A</sub>
32	GCATTAACACGCGCTAGCG GACG	Mutación W129A	IP <sub>5</sub> 2-K <sub>W129A</sub>
33	CGTCCGCTAGTGCTGTAA TGC	Mutación R45A	IP <sub>5</sub> 2-K <sub>R45A</sub>
34	GCATTAACAGCCGACTAGC GGACG	Mutación R45A	IP <sub>5</sub> 2-K <sub>R45A</sub>
35	GATTTCGTATACAGAAGGCT GCGAGGAATGATAAAGCAA TC	Mutación W129A/ R130A	IP <sub>5</sub> 2-K <sub>W129A/ R130A</sub>
36	GATTGCTTTATCATTCTCG CAGCCTTCTGTATACGAATC	Mutación W129A/ R130A	IP <sub>5</sub> 2-K <sub>W129A/ R130A</sub>
37	CGTCCGCTATGGGCTGTAA TGC	Mutación R130A	IP <sub>5</sub> 2-K <sub>R130A</sub>
38	GCAGGCGATAGCGCACAAT TACG	Mutación R130A	IP <sub>5</sub> 2-K <sub>R130A</sub>
39	CCCGGTGCTACCACCCCTA ATATGAGAGAACC	Mutación S253/S255	IP <sub>5</sub> 2-K <sub>S253/S255</sub>
40	GGGGGTGGTAGCACCGGGA GAACCAGCCCC	Mutación S253/S255	IP <sub>5</sub> 2-K <sub>S253/S255</sub>
41	CGGAATTCATGGGCGAGGA CGTGGGT	Introducción diana EcoRI	IP <sub>3</sub> 3-K <sub>158-201</sub>
42	CCCAAGCTTCTACGCCTTAA AACTCCCAAGTGTG	Introducción diana HindIII	IP <sub>3</sub> 3-K <sub>158-201</sub>
43	CTACGCCTGGGTGTAGCTGG CAGGGCAC	Introducción codon parada	IP <sub>3</sub> 3-K <sub>158-189</sub>
44	GTGCCCTGCCAGCTACACCC AGGCGTAG	Introducción codon parada	IP <sub>3</sub> 3-K <sub>158-189</sub>
45	GCCCTTTCAAGTAGCGCTAC GCCTG	Introducción codon parada	IP <sub>3</sub> 3-K <sub>158-183</sub>
46	CAGGCGTAGCGCTACTTGA AAGGGC	Introducción codon parada	IP <sub>3</sub> 3-K <sub>158-183</sub>

## 7.2.2 Cepas bacterianas

- **DH5α**

Cepa de *E. coli* que fue empleada para la propagación de plásmidos y clonaciones.

**Genotipo:** F<sup>+</sup> endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZ ΔM15  
Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ<sup>-</sup>.

**Referencia:** (Meselson y Yuan, 1968)

**Adquiridas a:** Invitrogen

- **Rosetta (DE3) pLysS**

Cepa de *E. coli* que fue empleada para la expresión de IP<sub>5</sub> 2-K y todos sus mutantes, para la expresión de IP<sub>3</sub> 3-K<sub>158-201</sub>, IP<sub>3</sub> 3-K<sub>158-189</sub> e IP<sub>3</sub> 3-K<sub>158-184</sub>.

**Genotipo:** F<sup>-</sup> ompT hsdS<sub>B</sub>(R<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) gal dcm λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) pLysSRARE (Cam<sup>R</sup>)

**Referencias:** (Wood, 1966) (Studier y Moffatt, 1986) (Moffatt y Studier, 1987) (Zahn, 1996)

**Adquiridas a:** EMD4 Biosciences/Millipore

- **BL21 (DE3)**

Cepa de *E. coli* que se empleó para la expresión de CaM.

**Genotipo:** F<sup>-</sup> ompT gal dcm lon hsdS<sub>B</sub>(r<sub>B</sub><sup>-</sup> m<sub>B</sub><sup>-</sup>) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])

**Referencias:** (Wood, 1966) (Studier y Moffatt, 1986)

**Adquiridas a:** EMD4 Biosciences/Millipore

- **C41 (DE3)**

Cepa de *E. coli* que se empleó para la expresión de IP<sub>3</sub> 3-K<sub>(128-461)</sub>.

**Genotipo:** F<sup>-</sup> ompT hsdSB (rB<sup>-</sup> mB<sup>-</sup>) gal dcm (DE3)

**Referencia:** (Miroux y Walker, 1996)

**Adquiridas a:** Lucigen

### 7.2.3 Vectores

- **pOPTG:** vector de expresión en bacterias realizado por la Dra. Olga Perisic (*Laboratory of Molecular Biology*, MRC, Cambridge, Reino Unido). Codifica para la Glutación S Transferasa (como etiqueta N-terminal) y presenta resistencia a ampicilina, empleando un promotor de tipo T7lac.
- **pKLSL<sub>t</sub>:** vector de expresión en bacterias realizado por el Dr José Miguel Mancheño (IQFR, CSIC, Madrid) (Angulo *et al.*, 2011). Codifica para el dominio de unión a azúcares de la lectina hemolítica LSL del hongo *L. sulphureus*, el cual se emplea como etiqueta de fusión en el extremo N-terminal. Presenta resistencia a kanamicina, y emplea como promotor una secuencia de tipo T7lac.

- **pET14b**: vector comercial de expresión en bacterias de la casa EMD4 Biosciences/Millipore.

#### 7.2.4 Reactivos, kits de biología molecular y enzimas

Los reactivos químicos empleados para purificación de proteínas fueron de calidad analítica, perteneciendo todos ellos a las marcas Sigma Aldrich, Fluka y Merck, salvo lo indicado expresamente.

Los componentes de origen microbiológico de los medios de cultivo utilizados fueron suministrados por Pronadisa/Conda.

Las enzimas de restricción utilizadas procedían de New England Biolabs. Los reactivos empleados en las reacciones de PCR fueron suministrados por New England Biolabs (polimerasa Taq) o EMD4 Biosciences/Millipore (polimerasa KOD). La DNA ligasa fue adquirida a Roche.

Los oligonucleótidos fueron adquiridos a Sigma Aldrich (oligonucleótidos 1-32) e IDT Technologies (oligonucleótidos 32-48).

Los marcadores de masa molecular de DNA fueron adquiridos a Bioline y los marcadores de masa molecular de proteínas fueron adquiridos a BioRad.

Los kits para biología molecular empleados fueron:

- **QIAquick® PCR Purification Kit** (Qiagen): se utilizó para la purificación de productos de PCR.
- **QIAquick® Gel Extraction Kit** (Qiagen): se utilizó para la purificación de fragmentos de DNA procedentes de electroforesis en geles de agarosa.
- **QIAprep® Spin Miniprep Kit** (Qiagen): se utilizó para la purificación de plásmidos procedentes de *E. coli*.

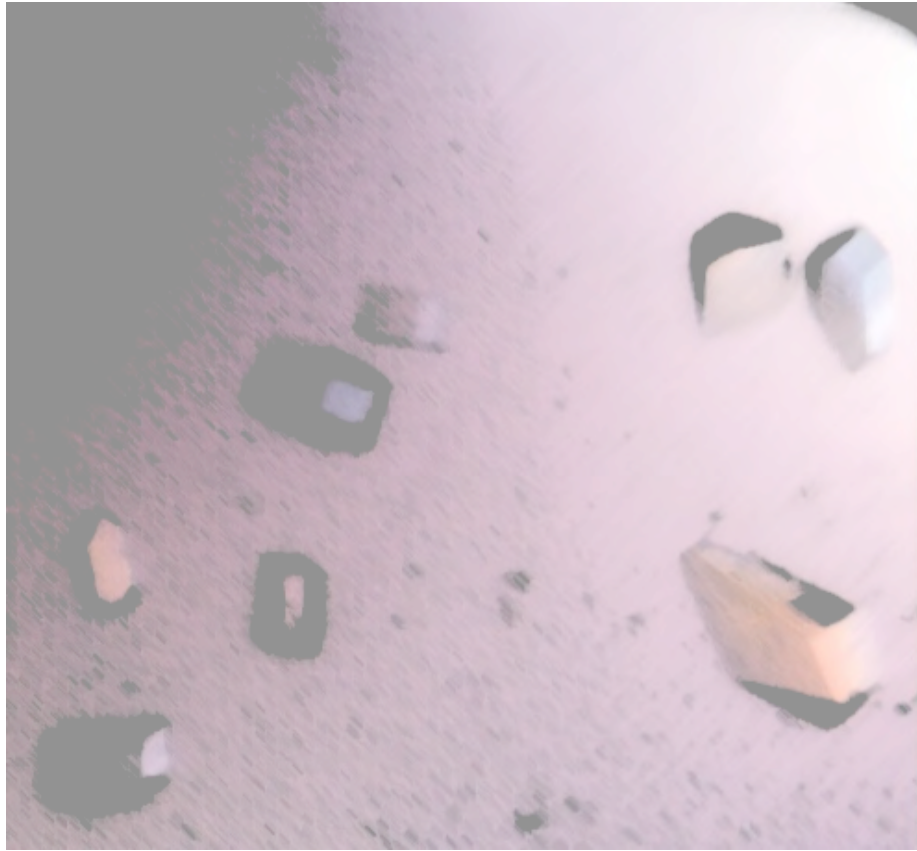
Las marcas del resto de productos empleados durante esta tesis están descritas en la sección Materiales y Métodos.

## 7.2.5 Resumen de las construcciones empleadas

Descripción	Fuente
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>R40V</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>R130I</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>K168A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>K168N</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>D368A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>D407A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>K411A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>N238A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>E85A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>C330A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>C333A</sub>	Este trabajo
pOPTG:: IP <sub>5</sub> 2-K <sub>H346Q</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K <sub>W129A</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K <sub>R45A</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K <sub>W129A/R130A</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K <sub>R130A</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>5</sub> 2-K <sub>S253/S255</sub>	Este trabajo
pOPTG:: IP <sub>3</sub> 3-K <sub>(128-461)</sub>	Dra. Olga Perisic ( <i>Laboratory of Molecular Biology</i> , MRC, Cambridge, Reino Unido)
pKLSL <sub>t</sub> :: IP <sub>3</sub> 3-K <sub>(158-202)</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>3</sub> 3-K <sub>(158-189)</sub>	Este trabajo
pKLSL <sub>t</sub> :: IP <sub>3</sub> 3-K <sub>(158-184)</sub>	Este trabajo
pET14b::CaM	Dra. Olga Perisic ( <i>Laboratory of Molecular Biology</i> , MRC, Cambridge, Reino Unido)

## 7.2.6 Resumen de las estructuras resueltas

Código PDB	Descripción
2axn	IP <sub>5</sub> 2-K de <i>A. thaliana</i> en complejo con IP <sub>6</sub> /AMPPNP
2axo	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>5</sub>
2axr	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>6</sub>
2axm	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>6</sub> /ADP
2axl	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>6</sub> /ADP/Pb <sup>2+</sup>
4axf	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>4</sub> /AMPNPP
4aqk	IP <sub>5</sub> 2-K <i>A. thaliana</i> en complejo con IP <sub>6</sub> /ADP (1 molécula en la unidad asimétrica)
4axc	IP <sub>5</sub> 2-K <i>A. thaliana</i> mutante W129A (IP <sub>5</sub> 2-K <sub>W129A</sub> )
4axd	IP <sub>5</sub> 2-K <sub>W129A</sub> en complejo con AMPNPP
4axe	IP <sub>5</sub> 2-K <sub>W129A</sub> en complejo con ADP
No depositada	CaM <i>H. sapiens</i>
No depositada	Complejo entre el dominio de unión a CaM de IP <sub>3</sub> 3-K A (residuos 158-183) humana y CaM humana



PUBLICACIONES



Jose Ignacio Baños-Sanz,<sup>a‡</sup>  
 Maider Villate,<sup>a‡</sup> Julia Sanz-  
 Aparicio,<sup>a</sup> Charles Alistair  
 Brearley<sup>b</sup> and Beatriz  
 González<sup>a\*</sup>

<sup>a</sup>Grupo de Cristalografía Macromolecular y  
 Biología Estructural, Instituto de Química-Física  
 'Rocasolano', CSIC, Serrano 119,  
 28006 Madrid, Spain, and <sup>b</sup>School of  
 Biological Sciences, University of East Anglia,  
 Norwich NR4 7TJ, England

‡ These authors contributed equally to this work  
 as first authors.

Correspondence e-mail: xbeatriz@iqfr.csic.es

Received 2 October 2009

Accepted 26 November 2009

## Crystallization and preliminary X-ray diffraction analysis of inositol 1,3,4,5,6-pentakisphosphate kinase from *Arabidopsis thaliana*

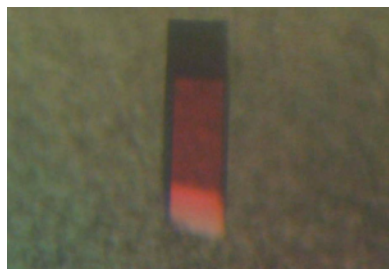
Inositol 1,3,4,5,6-pentakisphosphate kinase (IP<sub>5</sub> 2-K) is an enzyme involved in inositol metabolism that synthesizes IP<sub>6</sub> (inositol 1,2,3,4,5,6-hexakisphosphate) from inositol 1,3,4,5,6-pentakisphosphate (IP<sub>5</sub>) and ATP. IP<sub>6</sub> is the major phosphorus reserve in plants, while in mammals it is involved in multiple cellular events such as DNA editing and chromatin remodelling. In addition, IP<sub>6</sub> is the precursor of other highly phosphorylated inositols which also play highly relevant roles. IP<sub>5</sub> 2-K is the only enzyme that phosphorylates the 2-OH axial position of the inositide and understanding its molecular mechanism of substrate specificity is of great interest in cell biology. IP<sub>5</sub> 2-K from *Arabidopsis thaliana* has been expressed in *Escherichia coli* as two different fusion proteins and purified. Both protein preparations yielded crystals of different quality, always in the presence of IP<sub>6</sub>. The best crystals obtained for X-ray crystallographic analysis belonged to space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>, with unit-cell parameters *a* = 58.124, *b* = 113.591, *c* = 142.478 Å. Several diffraction data sets were collected for the native enzyme and two heavy-atom derivatives using a synchrotron source.

### 1. Introduction

Inositol phosphates (InsPs) are molecules that play important biological roles in various key cell metabolic and signalling processes. Their levels are strictly regulated by various inositol kinases (InsP Ks) and phosphatases. In the latter steps of InsP metabolism, highly phosphorylated InsPs such as inositol 1,2,3,4,5,6-hexakisphosphate (IP<sub>6</sub>; also known as phytic acid) are formed. IP<sub>6</sub> constitutes a major phosphorus reserve in plants and plays an essential role in processes as diverse as mRNA export (York *et al.*, 1999), DNA repair (Hanakahi & West, 2002), maintenance of basal resistance to plant pathogens (Murphy *et al.*, 2008), apoptosis (Agarwal *et al.*, 2009) and regulation of chromatin structure (Shen *et al.*, 2003; Steger *et al.*, 2003). In addition, IP<sub>6</sub> is the precursor of pyrophosphate inositols IP<sub>7</sub> and IP<sub>8</sub> (Mulugu *et al.*, 2007). Finally, several experiments have shown that IP<sub>6</sub> is a potential antitumour agent (Diallo *et al.*, 2008; Bozsik *et al.*, 2007).

The family of enzymes that catalyze the phosphorylation of inositol 1,3,4,5,6-pentakisphosphate (IP<sub>5</sub>) to form IP<sub>6</sub> are known as IP<sub>5</sub> 2-kinases (IP<sub>5</sub> 2-Ks). IP<sub>5</sub> 2-Ks have been cloned and characterized from several sources (Verbsky *et al.*, 2002; Phillippy *et al.*, 1994; Ives *et al.*, 2000; Sweetman *et al.*, 2006; Sun *et al.*, 2007; Suzuki *et al.*, 2007). Disruption of the IP<sub>5</sub> 2-K gene (*ipk1*) yields nonviable murine embryos (Verbsky *et al.*, 2005). IP<sub>5</sub> 2-K is also involved in the establishment of left–right organ asymmetry in mammals (Sarmah *et al.*, 2005). These enzymes present low sequence identity from yeast to mammals (about 11%) and do not present sequence homology to any other family of enzymes, including other families of inositol kinases. IP<sub>5</sub> 2-K is the only InsP K that phosphorylates the 2-axial position of the *myo*-inositide ring, whereas the other enzymes act on equatorial positions. For this reason, IP<sub>5</sub> 2-K represents an intriguing key point in cell biology, the understanding of which is crucial to obtaining insights into the regulation of inositol metabolism.

To date, the structures of several inositol kinases involved in soluble inositol regulation have been solved (González *et al.*, 2004;



© 2010 International Union of Crystallography  
 All rights reserved

Miller & Hurley, 2004; Holmes & Jøgl, 2006; Miller *et al.*, 2005). However, the structure of the IP<sub>5</sub> 2-K family remains unknown. Structural analysis of IP<sub>5</sub> 2-K will yield highly valuable information to fully understand the molecular mechanism of the substrate specificity of this peculiar family of InsP Ks that are essential for the functioning of mammals. In this study, we describe the expression, purification, crystallization and preliminary X-ray crystallographic analysis of IP<sub>5</sub> 2-K from *Arabidopsis thaliana*.

## 2. Experimental

### 2.1. Expression in *Escherichia coli*

The *ipk1* gene encoding the IP<sub>5</sub> 2-K protein from *A. thaliana* was amplified from At5G 42810 cDNA (Sweetman *et al.*, 2006) by polymerase chain reaction. The PCR products were cloned into two different vectors encoding different fusion proteins. The forward primer 5'-CGAGCTCATGGAGATGATTTTGGAGG-3' and the reverse primer 5'-CCGCTCGAGTTAGCTGTGGGAAGGTTTGAAGTTGC-3' were used to clone the PCR product into pKLSL<sub>1</sub> vector (Mancheño Gómez & Angulo Herrera, 2009). The forward primer 5'-GGAGATATACATATGGAGATGATTTTGGAGGAG-3' and the reverse primer 5'-GGGGTACCTTAGCTGTGGGAAGGTTTGAAGTTGC-3' were used to clone the PCR product into pOPTG vector. The primers include restriction sites for *Sac*I, *Xho*I, *Nde*I and *Kpn*I (shown in bold), respectively, for cloning the PCR product into the above vectors. The resulting proteins are fused with glutathione *S*-transferase (GST) and *Laetiporus sulphureus* lectin (LSL<sub>1</sub>) tags, respectively, and were expressed in *E. coli* Rosetta (DE3) pLysS strain. The bacteria were grown at 310 K until they reached an OD<sub>600</sub> of 1 and expression was induced by adding 0.3 mM IPTG and shaking for 15 h at 289 K.

### 2.2. Purification of the fusion proteins

For the GST-fusion protein, the cells were resuspended in buffer A (20 mM Tris-HCl pH 8, 50 mM NaCl, 2 mM DTT) plus protease inhibitors (complete EDTA-free tablets from Roche) and disrupted with a French press. After addition of 0.1% Triton X-100, the lysate was applied onto a heparin column, washed with buffer A and eluted with a 0–1 M NaCl gradient. The protein was loaded onto a GST-trap

HP column (GE Healthcare) and washed with buffer B (20 mM Tris-HCl pH 8, 200 mM NaCl, 2 mM DTT). The GST tag was removed by on-column incubation with TEV protease at 277 K overnight (protease:protein mass ratio 1:80). Finally, the IP<sub>5</sub> 2-K cleaved from GST (sample I) was eluted, concentrated and applied onto a gel-filtration column (16/60 Superdex 200) equilibrated in buffer B.

For the LSL<sub>1</sub>-fusion protein, the bacteria were lysed in buffer C (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT). The lysate was applied onto a Sepharose CL-4B column washed with buffer C and the protein was eluted with buffer C plus 200 mM lactose. The sample was diluted threefold to reduce the salt concentration, loaded onto a heparin column washed with buffer A and eluted with a 0–1 M NaCl gradient. After elution from the heparin column, the fused protein was cleaved by TEV protease (protease:protein mass ratio 1:80), shaking the sample gently at 277 K overnight, and IP<sub>5</sub> 2-K (sample II) was further purified and separated from LSL<sub>1</sub> by gel filtration using similar conditions as for sample I. The purified fraction was finally passed through a small Sepharose CL-4B column equilibrated in buffer B to remove the residual fused protein.

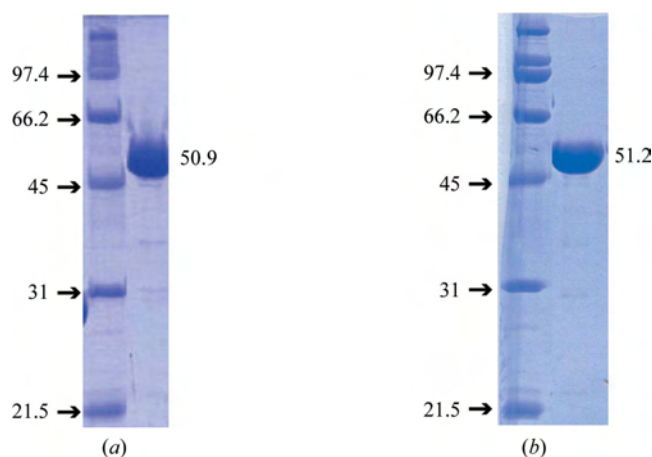
Both of the recombinant enzymes, which will be referred to as sample I and sample II, respectively, produced a well isolated peak in the gel-filtration chromatogram that could correspond to a monomer. The final yields of the purifications were 1 and 10 mg per litre of culture, respectively. The purity of the samples was confirmed by SDS-PAGE (Fig. 1). Both samples presented IP<sub>5</sub> 2-K activity, which was measured as described previously (Sweetman *et al.*, 2006). The final pure proteins were concentrated to around 10 mg ml<sup>-1</sup> and stored at 193 K.

### 2.3. Selenomethionine-substituted protein expression and purification

LSL<sub>1</sub>-IP<sub>5</sub> 2-K was expressed in minimal medium using the method reported by van Duyn *et al.* (1993). The selenomethionine-derivative protein was purified in a similar way to the native protein. The yield of the purified protein decreased dramatically in this protein form to a value of 0.6 mg per litre of culture.

### 2.4. Crystallization and cryoprotection

Initial crystallization conditions for protein samples I and II were investigated by high-throughput techniques with a NanoDrop robot (Innovadyne Technologies Inc.) using the commercial screen solutions Index and SaltRx from Hampton Research and PACT Suite and JCSG+ Suite from Qiagen. Duplicate crystallization assays were carried out for each protein sample with and without IP<sub>6</sub> using the sitting-drop vapour-diffusion method at 291 K in 96-well plates (Innovaplate SD-2 microplates, Innovadyne Technologies Inc.). In this method, drops consisting of 250 nl protein and 250 nl precipitant solution were mixed and equilibrated against 60 µl well solution. Many solutions containing PEG 3350 and PEG 6000 yielded crystals, but only in the assays using protein previously mixed with 2 mM IP<sub>6</sub>. Several strategies were used to optimize the crystallization conditions, which included adjusting the precipitant concentration and pH values, screening different additives (Additive Screen, Hampton Research) and streak-seeding. The final conditions were scaled up in 24-well plates (Cryscram plates, Hampton Research) using sitting-drop experiments by mixing 1 µl protein solution with 1 µl precipitant solution and equilibrating against 500 µl well solution. The optimization was approached independently for each protein form, including the selenomethionine-substituted protein. The best crystals were obtained from sample II mixed with 2 mM IP<sub>6</sub> and 2 mM ADP at 22% PEG 3350 and 0.1 M bis-tris pH 5.9. Several cryoprotectants



**Figure 1**  
SDS-PAGE analysis of IP<sub>5</sub> 2-K from *A. thaliana* for (a) sample I obtained from the GST-fusion protein and (b) sample II obtained from the LSL<sub>1</sub>-fusion protein. Both gels show the positions of the molecular-mass markers used; their molecular masses are indicated in kDa. Sample I and sample II have the same sequence except for the first N-terminal residues (GSH in sample I and GEFEL in sample II).

were tested, including glycerol; the most successful cryoprotection resulted from gradually increasing the PEG content to 35% of the crystal mother-liquor solution before flash-cooling to 100 K.

## 2.5. Formation of heavy-atom derivatives

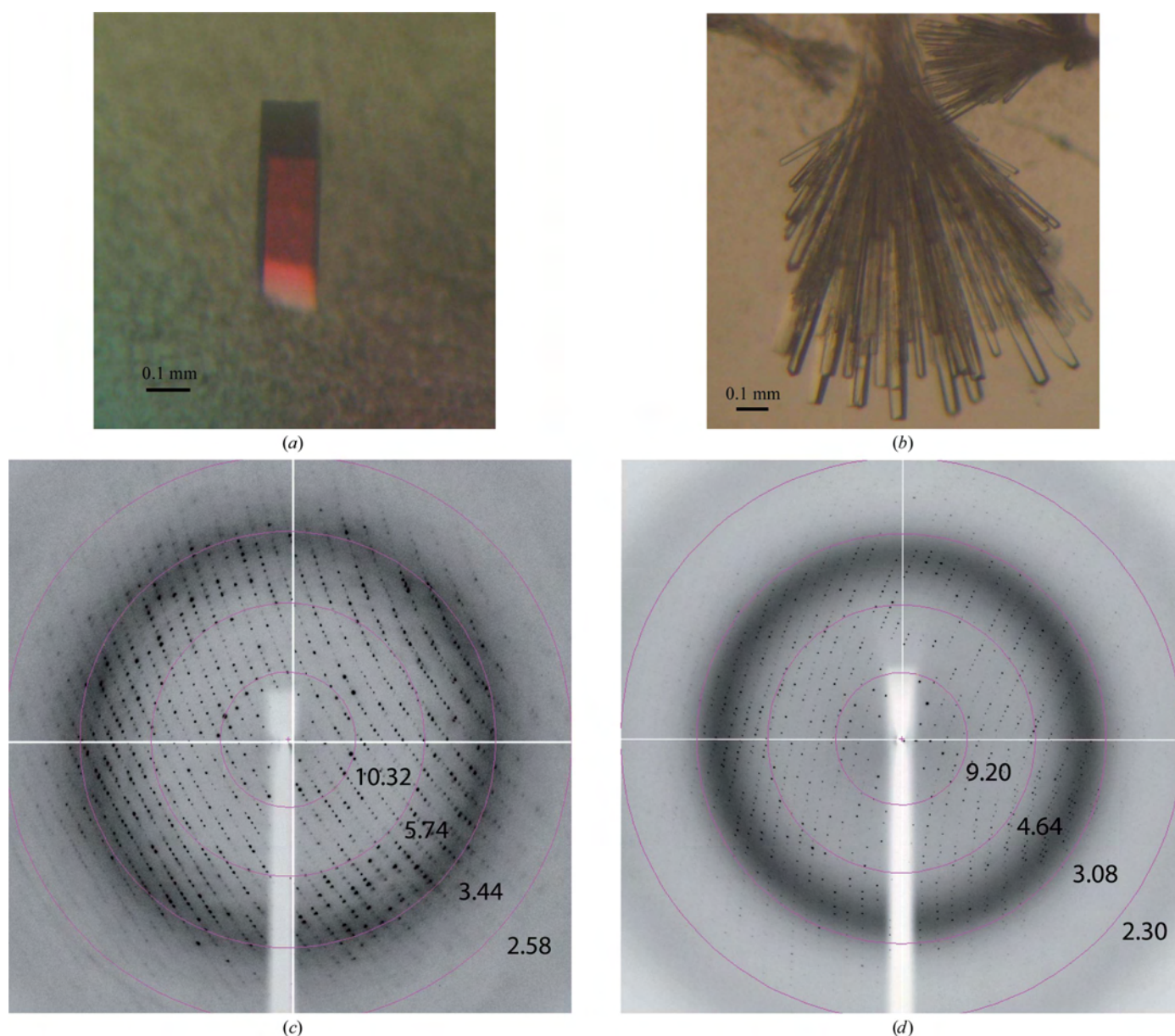
The best crystals (obtained in the presence of ADP and IP<sub>6</sub>) were subjected to soaks in various heavy-atom solutions. The heavy-atom compounds used were lead acetate and trimethyllead acetate (TMLA). For the soaks, these compounds were included in the cryoprotectant solution together with 2 mM IP<sub>6</sub> and 2 mM ADP and the crystals were left from 1 to 30 min in the heavy-atom high-PEG solution. The final soaking times and heavy-atom concentrations of the crystals used for X-ray diffraction analysis were 25 min and 100 mM in both cases.

## 2.6. Data collection and processing

Crystals were tested using in-house and synchrotron-radiation facilities. Native and MAD data sets were collected on various ESRF beamlines (Grenoble, France). The data sets were processed using the program *MOSFLM* (Leslie, 1990) and the *CCP4* suite (Collaborative Computational Project, Number 4, 1994). Heavy-atom positions were located using *SHELX* (Sheldrick, 2008) and were refined using *autoSHARP* (Vonrhein *et al.*, 2007). Crystal parameters and diffraction statistics are shown in Table 1.

## 3. Results and discussion

IP<sub>5</sub> 2-K from *A. thaliana* has been expressed as two different fusion proteins that allowed protein purification to apparent homogeneity. Interestingly, a good method for producing a considerable amount of IP<sub>5</sub> 2-K has been obtained (the method used to produce sample II).



**Figure 2**

Rod-shaped crystals of recombinant IP<sub>5</sub> 2-K from (a) sample I (0.25 × 0.11 × 0.11 mm) and (b) sample II (0.4 × 0.05 × 0.03 mm). (c, d) X-ray diffraction patterns of the crystals in (a) and (b) obtained using a synchrotron source. Numbers show the resolutions (Å) of the rings.



**Table 1**Data-collection statistics for the IP<sub>5</sub> 2-K crystals.

Values in parentheses are for the highest resolution shell.

Data set	Native				Lead (peak)	
	Sample I + IP <sub>6</sub>	Sample II + IP <sub>6</sub>	Sample II + IP <sub>6</sub>	Sample II + IP <sub>6</sub> , ADP	Sample II + IP <sub>6</sub> , ADP	Sample II + IP <sub>6</sub> , ADP
Cryoprotectant†	Glycerol	Glycerol	PEG 3350	PEG 3350	PEG 3350 + TMLA	PEG 3350 + lead acetate
Wavelength (Å)	0.93300	0.93300	0.93430	0.93340	0.94845	0.94780
Source	ESRF	ESRF	ESRF	ESRF	ESRF	ESRF
Beamline	ID14-2	ID14-2	ID14-1	ID14-1	ID23-1	ID23-1
Space group	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>	<i>P</i> <sub>2</sub> <sub>1</sub> <sub>2</sub> <sub>1</sub>
Unit-cell parameters (Å)						
<i>a</i>	60.172	59.538	57.669	58.124	57.046	58.063
<i>b</i>	115.067	117.523	112.501	113.591	112.583	110.970
<i>c</i>	236.254	150.446	139.720	142.478	141.081	138.694
Resolution (Å)	64.55–2.90 (3.06–2.90)	74.53–3.20 (3.37–3.20)	112.42–3.10 (3.27–3.10)	71.25–2.30 (2.42–2.30)	112.51–3.20 (3.37–3.20)	138.27–3.20 (3.37–3.20)
Unique reflections	37268 (5319)	18140 (2597)	17181 (2436)	42769 (6186)	15659 (2227)	15449 (2201)
<i>R</i> <sub>merge</sub> ‡	14.2 (42.5)	15.6 (47.1)	12.5 (44.5)	12.9 (45.6)	17.1 (40.9)	16.9 (44.2)
Completeness (%)	99.8 (99.8)	100.0 (100.0)	100.0 (100.0)	99.9 (100.0)	100.0 (100.0)	100.0 (100.0)
Mean multiplicity	5.3 (4.9)	5.7 (5.9)	7.6 (7.9)	7.3 (7.3)	7.7 (7.8)	6.7 (6.7)
Mean <i>I</i> / $\sigma$ ( <i>I</i> )	11.1 (2.1)	12.7 (2.9)	12.0 (4.4)	11.3 (4.2)	11.2 (5.9)	12.1 (4.0)
Wilson <i>B</i> factor (Å <sup>2</sup> )	80.7	72.0	67.1	36.7	43.2	47.7

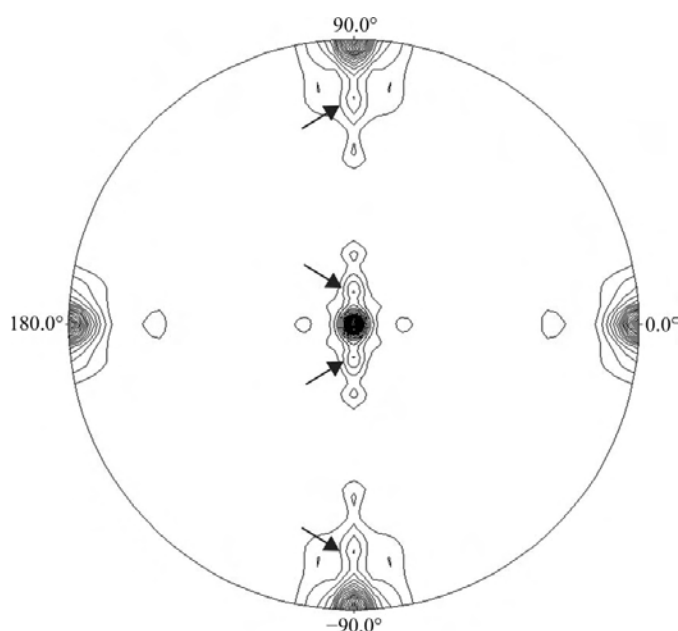
† The cryoprotectant includes the same products as were added to the sample. ‡  $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the *i*th observed amplitude of reflection *hkl* and  $\langle I(hkl) \rangle$  is the mean amplitude for measurements of reflection *hkl*.

Preliminary crystallization trials for both recombinant samples led to crystal clusters of thin plates or needles using PEG 3350 or PEG 6000 as precipitant. After refinement of several parameters, including a very fine tuning of the pH value, we obtained clusters of thicker plates or large rods; only a few assays with sample I led to isolated thick rods (Fig. 2). The best plate-shaped crystals from both samples appeared using 20% PEG 3350 and 0.1 *M* bis-tris propane pH 6.5. However, rods appeared at 17.5% PEG 3350 and 0.1 *M* sodium citrate pH 5.2 for sample I and 22% PEG 3350 and 0.1 *M* bis-tris pH 5.9 for sample II. The different protein-purification protocols and/or the remaining

N-terminal residues after tag removal could explain the different crystal properties of the two samples. We focused on the crystals that appeared as clusters of rods, since the X-ray diffraction data from the plate-shaped crystals could not be indexed. We initially used glycerol as a cryoprotectant and found that the sample I rods diffracted to 2.9 Å resolution (space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>, unit-cell parameters *a* = 60.175, *b* = 115.067, *c* = 236.254 Å; Table 1), while the thin sample II rods diffracted to 3.2 Å resolution (space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>, unit-cell parameters *a* = 59.538, *b* = 117.523, *c* = 150.448 Å; Table 1). Only the crystals from sample II were further optimized, as an analysis suggested that the sample I crystals were probably twinned. Cryo-protection was optimized by increasing the PEG concentration to 35% in the mother-liquor solution, which produced visibly lower damage to crystals than glycerol and a substantial change in the crystal unit cell (space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub>, unit-cell parameters *a* = 57.669, *b* = 112.501, *c* = 139.720 Å; Table 1). In addition, we observed that protein incubation with both products of the reaction, 2 mM ADP and 2 mM IP<sub>6</sub>, yielded crystals of higher quality, a few of which diffracted to 2.3 Å resolution (Table 1). In general, these crystals appeared in 4–5 days and continued growing over the following two weeks to maximum dimensions of 0.3 × 0.05 × 0.03 mm.

Assuming that two molecules are present in the asymmetric unit, the Matthews coefficient (Matthews, 1968) of crystals from sample II was 2.33 Å<sup>3</sup> Da<sup>−1</sup>, leading to a 47% solvent content within the unit cell. We have investigated the local symmetry relating the units in the asymmetric unit using the CCP4 package program POLARRFN (Kabsch, 1976). Several self-rotation functions were computed in the resolution range 15–3 Å, with Patterson vectors from 25 to 40 Å radius of integration. Analysis of self-rotation peaks revealed the presence of noncrystallographic twofold symmetry along the *bc* plane in the sample II crystals, although their height was rather low. The stereographic projection ( $\kappa = 180^\circ$  section) of the self-rotation is shown in Fig. 3.

We have expressed and purified selenomethionine-substituted protein, which yielded crystals under very similar conditions to the native crystals but that always grew as plates. The crystals were not suitable for X-ray data analysis since they were twinned. Therefore, we set up a series of our best crystallization assays with sample II in

**Figure 3**

Plot of the self-rotation function of IP<sub>5</sub> 2-K crystals from sample II using data between 15.0 and 2.2 Å resolution and a 25.0 Å radius of integration in the  $\kappa = 180^\circ$  section. The view is down the *c* axis.  $\varphi = 0^\circ$  and  $\varphi = 90^\circ$  correspond to the *a* and the *b* axes, respectively. The peaks highlighted show a noncrystallographic twofold axis along the *bc* plane, which is supported by the heavy-atom positions.

the presence of 2 mM ADP and 2 mM IP<sub>6</sub> and the crystals obtained were used for the preparation of heavy-atom derivatives. MAD experiments were carried out on the ID23-1 beamline using two crystals, one soaked in TMLA and other in lead acetate, which diffracted to >3 Å resolution (the peak data sets are shown in Table 1). It was possible to find several lead positions in the asymmetric units for both derivatives, which were refined and used for phase calculations. The phasing procedure confirmed the presence of two molecules in the asymmetric unit. Model building in the experimental electron-density maps obtained and refinement against the higher resolution native data set are currently in progress.

We thank Olga Perisic and Jose Miguel Mancheño for providing us with pOPTG and pKLSL<sub>t</sub> vectors, respectively. We thank the ESRF staff for providing time for and assistance with data collection. JIB is supported by an FPU fellowship (AP2008-00916) from the Ministerio de Educación. BG is supported by a 'Ramon y Cajal' fellowship (RYC-2006-002701). This work was supported by grants from the Comunidad de Madrid-CSIC (CCG07-CSIC/GEN-2232 and CCG08-CSIC/GEN-3490), Ministerio de Ciencia e Innovación (BFU2008-02897/BMC) and Ministerio de Educación y Ciencia (RYC-2006-002701). CAB was supported by the Biotechnology and Biological Sciences Research Council of the UK (grant BB/C514090/1).

## References

- Agarwal, R., Mumtaz, H. & Ali, N. (2009). *Mol. Cell. Biochem.* **328**, 155–165.
- Bozsik, A., Kokeny, S. & Olah, E. (2007). *Cancer Genomics Proteomics*, **4**, 43–52.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Diallo, J. S., Betton, B., Parent, N., Peant, B., Lessard, L., Le Page, C., Bertrand, R., Mes-Masson, A. M. & Saad, F. (2008). *Br. J. Cancer*, **99**, 1613–1622.
- González, B., Schell, M. J., Letcher, A. J., Veprintsev, D. B., Irvine, R. F. & Williams, R. L. (2004). *Mol. Cell*, **15**, 689–701.
- Hanakahi, L. A. & West, S. C. (2002). *EMBO J.* **21**, 2038–2044.
- Holmes, W. & Jorgl, G. (2006). *J. Biol. Chem.* **281**, 38109–38116.
- Ives, E. B., Nichols, J., Wente, S. R. & York, J. D. (2000). *J. Biol. Chem.* **275**, 36575–36583.
- Kabsch, W. (1976). *Acta Cryst.* **A32**, 922–923.
- Leslie, A. G. W. (1990). *Crystallographic Computing 5. From Chemistry to Biology*, edited by D. Moras, A. D. Podjarny & J.-C. Thierry, pp. 50–61. Oxford University Press.
- Mancheño Gómez, J. M. & Angulo Herrera, I. (2009). Patent WO/2009/121994.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Miller, G. J. & Hurley, J. H. (2004). *Mol. Cell*, **15**, 703–711.
- Miller, G. J., Wilson, M. P., Majerus, P. W. & Hurley, J. H. (2005). *Mol. Cell*, **18**, 201–212.
- Mulugu, S., Bai, W., Fridy, P. C., Bastidas, R. J., Otto, J. C., Dollins, D. E., Haystead, T. A., Ribeiro, A. A. & York, J. D. (2007). *Science*, **316**, 106–109.
- Murphy, A. M., Otto, B., Brearley, C. A., Carr, J. P. & Hanke, D. E. (2008). *Plant J.* **56**, 638–652.
- Phillippy, B. Q., Ullah, A. H. & Ehrlich, K. C. (1994). *J. Biol. Chem.* **269**, 28393–28399.
- Sarmah, B., Latimer, A. J., Appel, B. & Wente, S. R. (2005). *Dev. Cell*, **9**, 133–145.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Shen, X., Xiao, H., Ranallo, R., Wu, W. H. & Wu, C. (2003). *Science*, **299**, 112–114.
- Steger, D. J., Haswell, E. S., Miller, A. L., Wente, S. R. & O'Shea, E. K. (2003). *Science*, **299**, 114–116.
- Sun, Y., Thompson, M., Lin, G., Butler, H., Gao, Z., Thornburgh, S., Yau, K., Smith, D. A. & Shukla, V. K. (2007). *Plant Physiol.* **144**, 1278–1291.
- Suzuki, M., Tanaka, K., Kuwano, M. & Yoshida, K. T. (2007). *Gene*, **405**, 55–64.
- Sweetman, D., Johnson, S., Caddick, S. E., Hanke, D. E. & Brearley, C. A. (2006). *Biochem. J.* **394**, 95–103.
- Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). *J. Mol. Biol.* **229**, 105–124.
- Verbsky, J., Lavine, K. & Majerus, P. W. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8448–8453.
- Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W. & Wente, S. R. (2002). *J. Biol. Chem.* **277**, 31857–31862.
- Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
- York, J. D., Odom, A. R., Murphy, R., Ives, E. B. & Wente, S. R. (1999). *Science*, **285**, 96–100.

# Inositol 1,3,4,5,6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide binding site for axial 2-OH recognition

Beatriz González<sup>a,1</sup>, Jose Ignacio Baños-Sanz<sup>a</sup>, Maider Villate<sup>a</sup>, Charles Alistair Brearley<sup>b</sup>, and Julia Sanz-Aparicio<sup>a</sup>

<sup>a</sup>Grupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física Rocasolano, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain; and <sup>b</sup>School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

Edited by Susan S. Taylor, University of California at San Diego, La Jolla, CA, and approved March 24, 2010 (received for review November 10, 2009)

Inositol phosphates (InsPs) are signaling molecules with multiple roles in cells. In particular Ins(1,2,3,4,5,6)P<sub>6</sub> (InsP<sub>6</sub>) is involved in mRNA export and editing or chromatin remodeling among other events. InsP<sub>6</sub> accumulates as mixed salts (phytate) in storage tissues of plants and plays a key role in their physiology. Human diets that are exclusively grain-based provide an excess of InsP<sub>6</sub> that, through chelation of metal ions, may have a detrimental effect on human health. Ins(1,3,4,5,6)P<sub>5</sub> 2-kinase (InsP<sub>5</sub> 2-kinase or Ipk1) catalyses the synthesis of InsP<sub>6</sub> from InsP<sub>5</sub> and ATP, and is the only enzyme that transfers a phosphate group to the axial 2-OH of the myo-inositide. We present the first structure for an InsP<sub>5</sub> 2-kinase in complex with both substrates and products. This enzyme presents a singular structural region for inositide binding that encompasses almost half of the protein. The key residues in substrate binding are identified, with Asp368 being responsible for recognition of the axial 2-OH. This study sheds light on the unique molecular mechanism for the synthesis of the precursor of inositol pyrophosphates.

InsP<sub>5</sub> 2-kinase | IP<sub>5</sub> | IP<sub>6</sub> | ipk1 | crystal structure

Inositol phosphates (InsPs) are soluble signaling molecules that have diverse and important biological roles (1, 2). Inositol 1,2,3,4,5,6-hexakisphosphate (InsP<sub>6</sub> or IP<sub>6</sub>), known as phytic acid, plays an essential role in cell biology. For example, IP<sub>6</sub> is involved in processes as diverse as mRNA export (3), DNA repair (4), maintenance of basal resistance to plant pathogens (5), apoptosis (6), and regulation of chromatin structure (7, 8). In addition, IP<sub>6</sub> is the precursor of pyrophosphate inositols IP<sub>7</sub> and IP<sub>8</sub> (9, 10). A primary aspect of IP<sub>6</sub> is the key role it plays in plant biology, reflected in considerable commercial interest in this compound (11, 12). IP<sub>6</sub> constitutes the major phosphorus reserve in plant seeds, being a source not only of phosphorus, but also of inositol and minerals during development and germination (12). Concerns arise from the fact that IP<sub>6</sub> can present detrimental effects on human health and on the environment (reviewed in ref. 13). Grain-based diets provide high levels of IP<sub>6</sub>, a potent chelator of essential ions as Zn<sup>2+</sup> and Fe<sup>2+</sup>, causing deficiencies that contribute to malnutrition. Monogastric animals are unable to digest IP<sub>6</sub>, which is excreted as salts that contribute to eutrophication of waterways. There is, consequently, considerable interest in the manipulation of plant phytate, largely by genetic approaches, though the desirability of this is not without dispute (11–15). Indeed, there are several lines of evidence that show benefits of IP<sub>6</sub> as antioxidant or a potential antitumor agent (16, 17). The IP<sub>6</sub> precursor, inositol 1,3,4,5,6-pentakisphosphate (InsP<sub>5</sub> or IP<sub>5</sub>), also presents a role in induction of apoptosis (18).

The family of enzymes that catalyzes the phosphorylation of IP<sub>5</sub> to form IP<sub>6</sub> are known as Ipk1, after the ortholog first identified in yeast (3) or InsP<sub>5</sub> 2-kinases (IP<sub>5</sub> 2-Ks), after the enzyme activity characterized from various sources (19–24). It is only recently, however, that important physiological roles other than a biosynthetic enzyme have been ascribed to IP<sub>5</sub> 2-K. Thus, disruption of the IP<sub>5</sub> 2-K gene (*Ipk1*) yields non viable murine embryos (25). IP<sub>5</sub> 2-K is also involved in the establishment of left–right

organ asymmetry in mammals (26). As a family, IP<sub>5</sub> 2-Ks present very low sequence homology from yeast to mammals, with only a few sequence motifs being conserved. IP<sub>5</sub> 2-Ks are unique among inositol phosphate kinases (InsP Ks) in that they phosphorylate the axial 2-position of the myo-inositide ring, whereas the other enzymes act on equatorial positions of the ring.

The first structure of an InsP K became available in 2004 (27, 28), from IP<sub>3</sub> 3-Kinase (IP<sub>3</sub> 3-K), a family not represented in yeast or plant genomes. Based on sequence similarity, the InsP K enzymes encompass three different structural families. The first family is formed by the inositol polyphosphate kinases (IPKs), to which IP<sub>3</sub> 3-Ks belong, and this family displays some degree of structural homology with protein kinases (PKs) (27). Another family of InsP Ks have been shown to display the ATP-grasp fold (10, 29). A third structural family is formed by IP<sub>5</sub> 2-Ks, their structure remaining unknown, although Cheek et al. (30) have postulated that IP<sub>5</sub> 2-K could display some structural similarity to IPKs.

Herein we present the crystal structure of *Arabidopsis thaliana* IP<sub>5</sub> 2-K (*At*-IP<sub>5</sub> 2-K, *At*Ipk1), as well as a detailed description of the substrate and product interactions that have allowed us an in-depth characterization of this family fold and active site. The structure of IP<sub>5</sub> 2-K demonstrates membership, at a distance, of the IPK family of proteins. However, the IP<sub>5</sub> 2-K structure shows a unique and large inositide-binding pocket, which explains how the myo-inositol 2-OH axial position that accepts the phosphate is recognised. In addition, IP<sub>5</sub> 2-K structure reveals key features of IPKs that elucidate the functionality of this essential family of enzymes.

## Results

**IP<sub>5</sub> 2-K Structure.** The structure of *At*-IP<sub>5</sub> 2-K has been solved by multiwavelength anomalous dispersion (MAD) techniques and the structures of different IP<sub>5</sub> 2-K crystal complexes have been obtained: two binary complexes (IP<sub>5</sub> 2-K/IP<sub>5</sub> and IP<sub>5</sub> 2-K/IP<sub>6</sub>), and two ternary complexes (IP<sub>5</sub> 2-K/IP<sub>5</sub>/AMPPNP and IP<sub>5</sub> 2-K/IP<sub>6</sub>/ADP) (Table S1 and Fig. S1). IP<sub>5</sub> 2-K fold consists of two lobes connected by a hinge, which we will refer as N-terminal lobe (residues 1–152) and C-terminal lobe (161–432) (Fig. 1 A and B). Within the C-lobe there are three noncontiguous regions folded in a unique assembly involved in the inositide binding (C<sub>IP</sub>-lobe).

Author contributions: B.G., C.A.B., and J.S.-A. designed research; B.G., J.I.B.-S., and M.V. performed research; B.G., J.I.B.-S., M.V., C.A.B., and J.S.-A. analyzed data; and B.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The structural coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2xal, 2xam, 2xan, 2xao and 2xar).

To whom correspondence should be addressed. E-mail: xbeatriz@iqfr.csic.es.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.0912979107/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.0912979107/-DCSupplemental).



The N-terminal lobe (N-lobe) has an  $\alpha + \beta$  fold (Fig. 1*B* and Fig. S2*A*), the core of this lobe being formed by an antiparallel and twisted main  $\beta$ -sheet comprised of five  $\beta$ -strands in 1-2-3-6-4 order. Strands  $\beta$ 1– $\beta$ 2 are connected by a glycine rich loop that is involved in nucleotide binding. Strands  $\beta$ 3– $\beta$ 4 and  $\beta$ 4– $\beta$ 6 are connected through long regions built up mainly of  $\alpha$ -helices, named as “N-I” and “N-II,” respectively. N-I is comprised of a mobile loop and three antiparallel helices ( $\alpha$ 2– $\alpha$ 4). Finally, N-II starts with a helix ( $\alpha$ 5) that crosses the whole  $\beta$ -sheet, followed by a small and antiparallel helix and the short  $\beta$ 5.

The C-lobe is almost twice the size of the N-lobe and also has an  $\alpha + \beta$  fold (Fig. 1*B* and Fig. S2*B*). The  $\beta$ -strands form an antiparallel twisted  $\beta$ -sheet connected in order 9-8-7-10-12-11. The  $\alpha$ -helices in this lobe are distributed in three noncontinuous regions that emerge from the  $\beta$ -sheet and comprise almost half of the protein. Helices  $\alpha$ 11 and  $\alpha$ 12, preceded by the long loop L3, fold over the  $\beta$ -sheet. All other  $\alpha$ -helices form a well-defined assembly with a very irregular fold that we have named as  $C_{IP}$  lobe, because it constitutes the main inositide-binding site. The first region ( $C_{IP-I}$ ), connects strands  $\beta$ 7– $\beta$ 8 and is formed by  $\alpha$ 7– $\alpha$ 10.  $C_{IP-I}$  starts with a long loop (L1) followed by  $\alpha$ -helices that come back to the  $\beta$ -sheet in a zigzag fashion, ending at the loop L2. The second region ( $C_{IP-II}$ ) connects  $\alpha$ 12– $\beta$ 10 and is comprised of  $\alpha$ 13– $\alpha$ 16, which form a kind of coil joined to the  $\beta$ -sheet by the short L4. Finally,  $C_{IP-I}$  and  $C_{IP-II}$  are held together through L1, and the C-terminal helix. This helix and its preceding loop L5 constitute  $C_{IP-III}$ . The loops L1–L5, located at the junction of the C-lobe  $\beta$ -sheet and helices, are related to four IP<sub>5</sub> 2-K conserved sequence motifs (motif I–IV) (Fig. 1*C*). There are two additional motifs conserved along IP<sub>5</sub> 2-Ks, located in  $\alpha$ 8 and  $\alpha$ 13– $\alpha$ 14 of the  $C_{IP}$ -lobe as will be seen later.

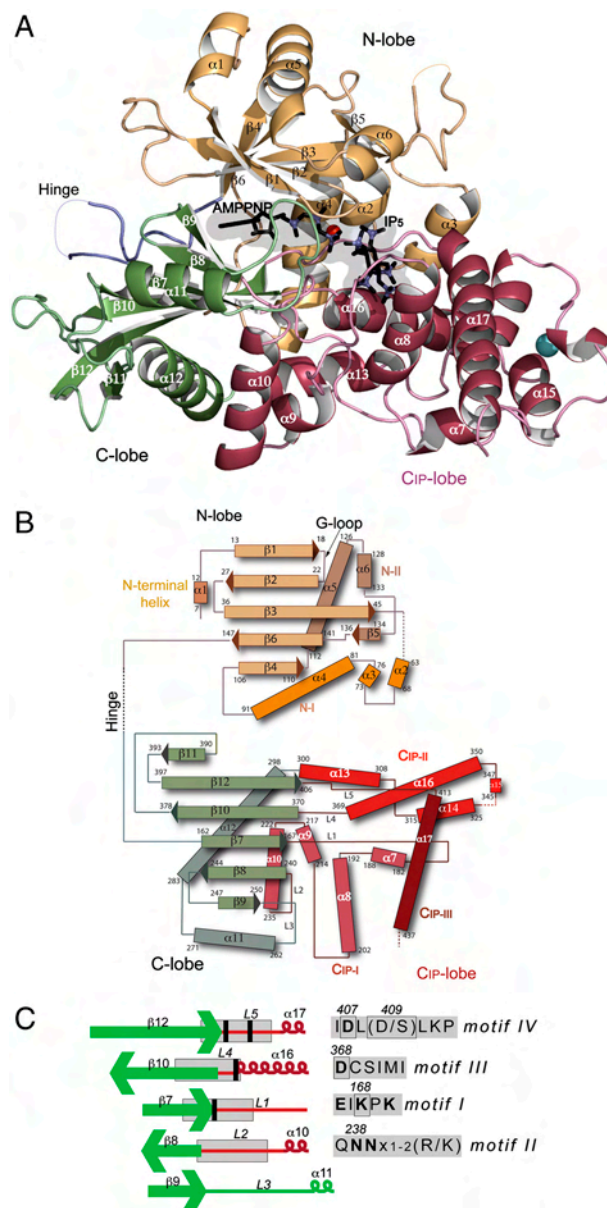
IP<sub>5</sub> 2-K crystallizes with two molecules in the asymmetric unit. There are only a few hydrogen bonds between the two molecules, with 1, 101 Å<sup>2</sup> buried surface. This suggests that it is not likely to be a physiological dimer. No significant differences are seen in the active sites of the two molecules.

### IP<sub>5</sub> 2-K Substrate and Product Recognition.

**Nucleotide-binding site.** The ATP analogue is bound between the N- and C-lobes, at the interface created by both  $\beta$ -sheets of the enzyme (Figs. 1*A* and 2*A*, and Fig. S3). The adenine N1 and N6 are linked to hinge residues (Asn147 and His149, respectively). Several residues from both lobes provide a hydrophobic environment for the adenine and ribose: Arg16, Val24, Val38, Leu146, His149, Met372, and Ile406. The oxygen atoms of the ribose are coordinated to Glu166. A weaker interaction is found between the ribose O3\* and Arg241. P $\alpha$  is coordinated to Arg40, whereas the P $\beta$  is coordinated to several residues from the G-loop (Asn22, and the main chain of Gly19, Gly20, Ala21, and Asn22) and Arg40. Finally, P $\gamma$  is recognised by the G-loop Gly20, Lys168, Asp407, and Ser409. Interactions with the latter two residues occur via Mg1, which in turn is coordinated to P $\beta$ . P $\gamma$  is further coordinated by Asp 368.

Interestingly, the conserved motifs I–IV are all involved in nucleotide recognition and in particular, residues located at L1, L4, and L5 surround the phosphoryl transference region (Fig. 2*D*).

**Inositide-binding site.** IP<sub>5</sub> 2-K substrate (IP<sub>5</sub>) is tightly bound to the enzyme through all phosphates groups (Fig. 2*B* and Fig. S3). P1 makes two hydrogen bonds with Arg130 from the N-lobe, whereas P3 is bound to both the N-lobe (Arg45) and the C-lobe (Lys411 and Arg415). Interestingly, P4, P5, and P6 are coordinated only by residues from the  $C_{IP}$ -lobe. P4 coordinates to Lys411, Arg415, and Tyr419, P5 coordinates to Lys170 (and to His196 and Arg192 through water molecules), and P6 coordinates to Lys170, Lys200, and Asn238. A water-mediated link is also produced with Asn238 and Asn239. Finally, the acceptor 2-OH interacts to Asp368, and Lys168. Asp368 seems to be



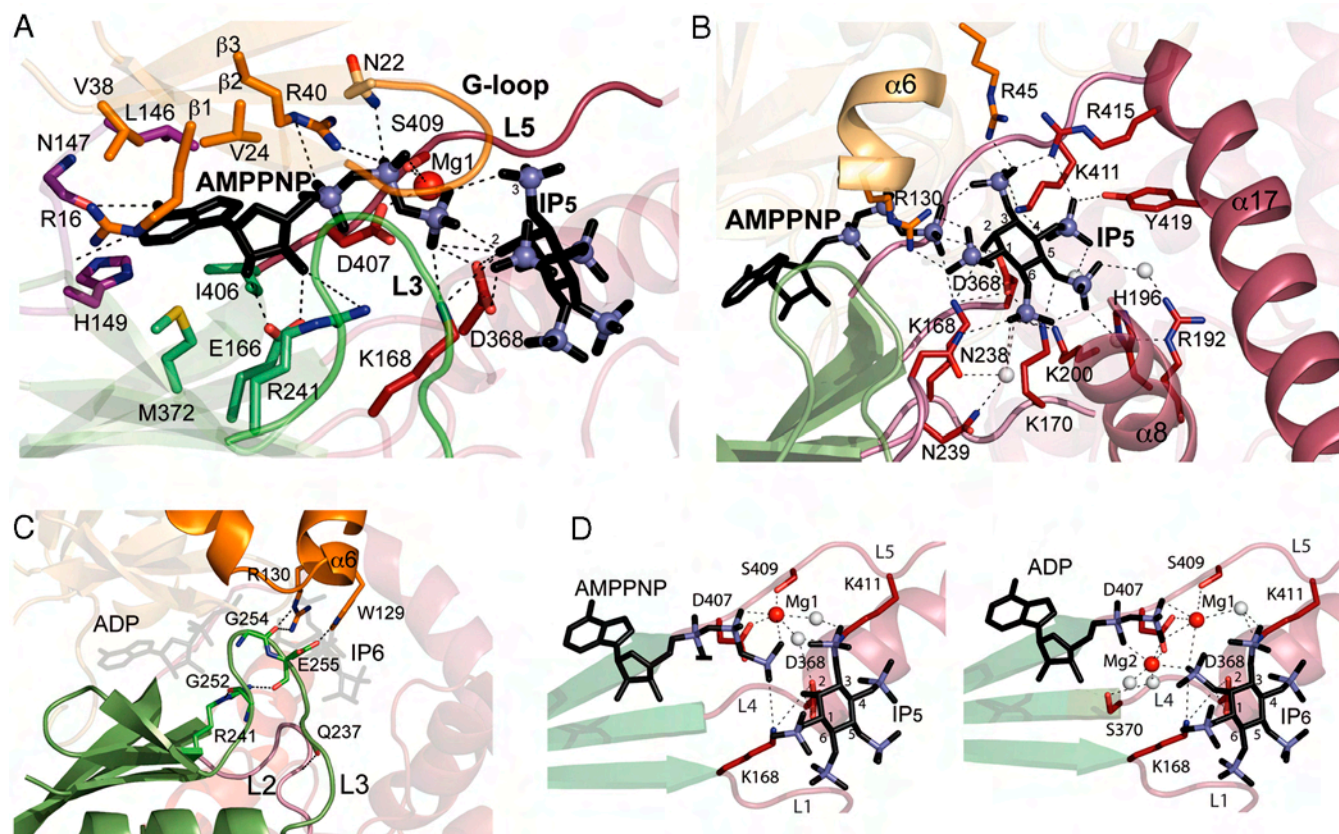
**Fig. 1.** The structure of IP<sub>5</sub> 2-K from *A. thaliana*. (A) Cartoon representation of *At*-IP<sub>5</sub> 2-K structure, in which the N-lobe is colored in light orange, the hinge in slate blue, the C-lobe  $\beta$ -sheet region in green, and the  $C_{IP}$ -lobe in raspberry. The substrates are shown as black sticks, with the phosphates as red spheres. The magnesium and zinc ions are shown as red and cyan spheres, respectively. (B) IP<sub>5</sub> 2-K topology diagram. The color code is similar, applying different orange shades for the N-lobe regions ( $\beta$ -sheet, N-terminal helix, N-I, and N-II) and different red shades for the three  $C_{IP}$  elements ( $C_{IP-I}$ ,  $C_{IP-II}$ , and  $C_{IP-III}$ ). (C) Scheme of the C-lobe  $\beta$ -sheet and  $\alpha$ -helices interface showing loops L1 to L5. The gray shade squares show the position of four IP<sub>5</sub> 2-K's conserved motifs. Residues invariant along the species are in bold letter. Relevant residues for catalysis are squared and its position is represented as black rods.

essential in recognizing of the inositide 2-OH, likely conferring specificity for the substrate conformer.

In conclusion, the  $C_{IP}$ -lobe provides residues for binding all IP<sub>6</sub> phosphates, except P1, coming from motifs I–IV, motif <sup>92</sup>RxxMHQxLK. The N-lobe helps to inositide binding by coordinating P1 and P3. The strong interaction that all IP<sub>5</sub> phosphates establish with the protein explains why other less substituted InsPs lacking phosphates on 1-, 3-, or 5- or 3,5-positions are substrates of IP<sub>5</sub> 2-K (20, 24, 31).

Finally, L3 seems to constrict the inositide-binding site by making interactions with  $\alpha$ 6 from N-lobe and motif II from





**Fig. 2.** IP<sub>5</sub> 2-K active site. AMPPNP (A) and IP<sub>5</sub> (B) interactions as found in the ternary substrates complex. Residues involved in AMPPNP or inositol recognition are shown in sticks. Water molecules are represented in white and hydrogen bonds as dashed lines. (C) L3 atomic interactions as observed in the ternary products complex. This loop interacts with the side chain of Arg130 and Trp129 from α6, and residues in or near L2 as Gln237 and Arg241. L3 keeps together regions involved in inositol recognition. (D) Substrates (Left) and products (Right) recognition and metal interactions in both ternary complexes. The substrates complex shows only one metal, Mg1, bound to the L5 residues Asp407 and Ser409, whereas the products complex shows two metals, Mg1 and Mg2, both bound to Asp407 and bridging inositol and nucleotide ligands.

the C<sub>IP</sub>-lobe (Fig. 2C). Interactions between L3 and G-loop through water molecules are also observed. We suspect that *At*-IP<sub>5</sub> 2-K undergoes a conformational change in presence or absence of the inositol (for example, it only crystallizes in presence of inositol), which could be attributed to changes in L3.

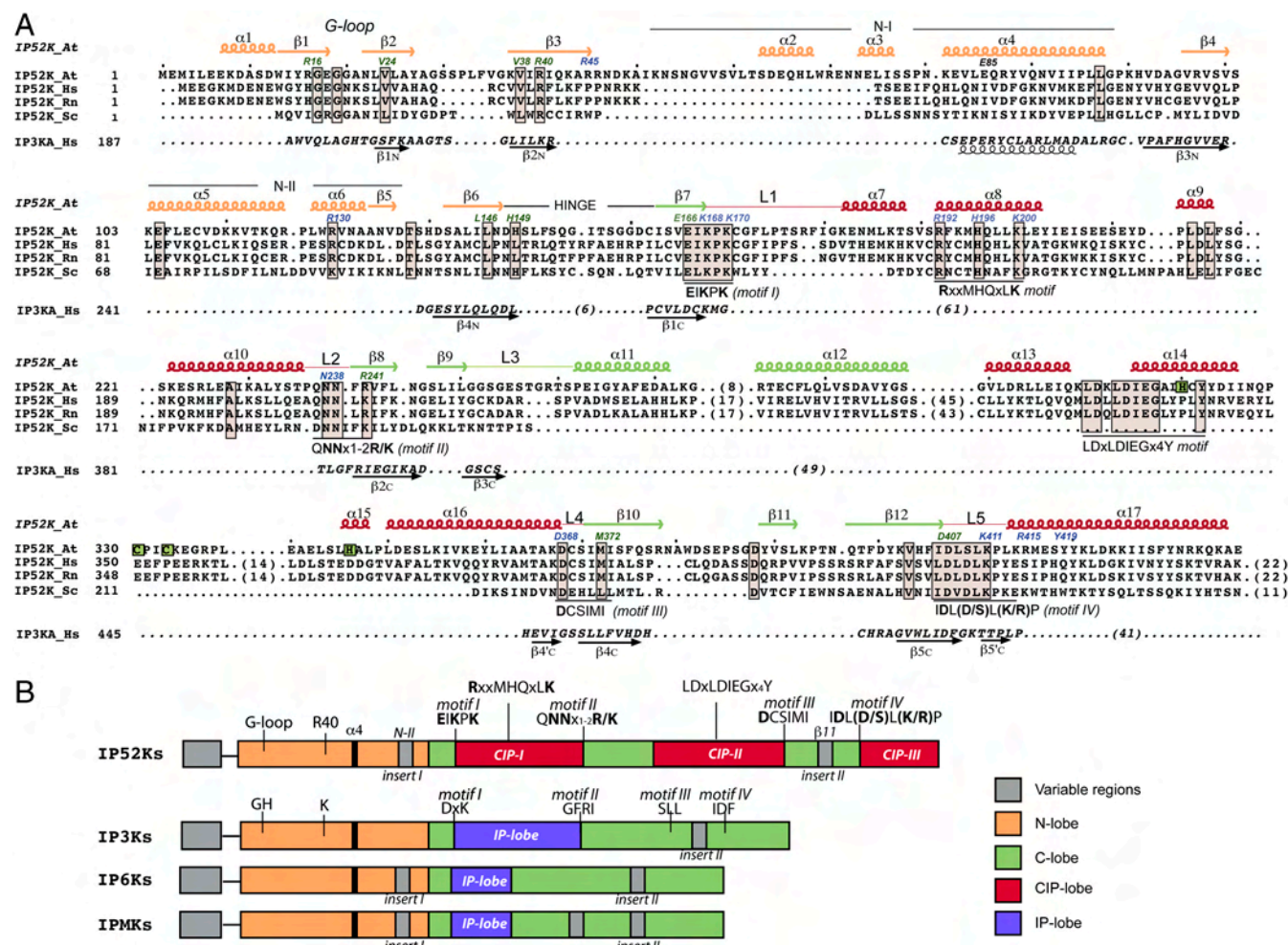
**Different ligand recognition steps and metal binding.** The complexes obtained do not display great structural differences (maximal rmsd of 0.479 Å), the inositol interactions being essentially conserved. However, magnesium binding is only observed in ternary complexes with both substrates or products (Fig. 2D and Fig. S4A). In the complex with both substrates, the inositol P3 and 2-OH interact with P<sub>γ</sub>, whereas Mg1 orients this P<sub>γ</sub>. In the complex with products, two metals (Mg1 and Mg2), are bridging the inositol P2 to the nucleotide P<sub>α</sub> and P<sub>β</sub>. Detailed interactions of metal- and ligand- protein recognition are shown in Fig. 2D. The lead derivative crystal, obtained from a ternary ADP/IP<sub>6</sub> complex, showed two Pb<sup>2+</sup> ions in similar positions to both Mg<sup>2+</sup> ions, mimicking this cofactor.

There are no great structural changes among complexes upon nucleotide binding. A remarkable feature deduced from the binary complexes is that the inositol apparently does not leave room for the nucleotide to enter (Fig. S4B), suggesting that either L3 moves from a putative closed to an opened conformation allowing nucleotide entrance, or the nucleotide must bind first into the active site. It remains to be determined whether IP<sub>5</sub> 2-K presents an ordered-ligand kinetics.

**IP<sub>5</sub> 2-Ks from Plants Conserve a Zinc-Binding Motif.** *At*-IP<sub>5</sub> 2-K presents a Zn motif in the C<sub>IP</sub>-II (α14–α15). The electron density

map shows a high density peak connected to His320, Cys330, Cys333, and His346 (Fig. S5A). This site has been attributed to a Zn<sup>2+</sup> cation, although *At*-IP<sub>5</sub> 2-K was not previously identified as a zinc-binding enzyme. Inductively coupled plasma-optical emission spectroscopy (ICP-OES) analysis of *At*-IP<sub>5</sub> 2-K confirmed the presence of 1 mol of zinc per mol of protein. All IP<sub>5</sub> 2-Ks from plants conserve all the four Zn-binding residues (Fig. S5B). We have performed site-directed mutagenesis on the Zn-binding site, obtaining that C330S, C333S, and H346N single mutants are not stable under native conditions (Fig. S5C). In addition the incremental addition of EDTA to the wild-type protein reduced intrinsic tryptophan fluorescence by approximately 20% in an irreversible process (Fig. S5D). However incubation of the wild-type with EDTA has no effect on enzyme activity (Fig. S5E). Altogether, these observations suggest that Zn has a structural role in the proper folding of protein, but it seems not essential for enzymatic activity.

**IP<sub>5</sub> 2-Ks Family and Mutagenesis Analysis.** IP<sub>5</sub> 2-Ks from different sources have low sequence conservation (11%), *A. thaliana* and human enzymes being 40% homologous. However, small sequence motifs tend to be conserved in all IP<sub>5</sub> 2-Ks. We have performed a structural alignment among the 67 different IP<sub>5</sub> 2-Ks yielded by PFAM (<http://pfam.sanger.ac.uk/>), representatives being shown in Fig. 3A. The β-sheets are conserved in all species. The N-lobe helices are conserved, except α2 and α3 in region N-I, that seem specific for the plant enzymes, and conform a flexible region partially disordered in the crystal, with no apparent implications in substrate binding or catalysis. In the C-lobe, mammalian enzymes conserve all the α-helices described in *At*-IP<sub>5</sub> 2-K



**Fig. 3.** The IP<sub>5</sub> 2-K family. (A) Structural alignment of At-IP<sub>5</sub> 2-K with IP<sub>5</sub> 2-Ks from other sources (Hs:human, Rn:rat, and Sc:yeast). Secondary elements are shown and named on top, and colored in the same code. Zn binding residues from At-IP<sub>5</sub> 2-K are inside green squares. Conserved residues are gray squared and motifs are shown below. Residues shown in dark green and blue are ATP and inositol ligands, respectively. The IP<sub>3</sub> 3-K structural elements conserved with IP<sub>5</sub> 2-K and their sequence are shown below. (B) Schematic domain representation of IP<sub>5</sub> 2-Ks vs. the IPK family (IP<sub>3</sub> 3-K, IP<sub>6</sub> Ks and IPm Ks). The black rod represents  $\alpha 4$ , which is conserved in all IPKs, and is similar to PKs  $\alpha C$ . The IP<sub>5</sub> 2-K G-loop, R40, and conserved motifs are marked on top, and the equivalent IPK regions are shown.

showing two additional main insertions at C<sub>IP-II</sub>. The first 45 residues insertion is found between  $\alpha 12$  and  $\alpha 13$ , with no secondary structure predicted. A second insertion, 14 residues long and between  $\alpha 14$  and  $\alpha 15$ , is predicted to form an additional helix in mammals. This last insertion is located in the Zn-binding motif found in At-IP<sub>5</sub> 2-K, and interestingly the Zn ligands are not conserved in mammals or in other species. The yeast and fungi enzymes present a shorter C<sub>IP-lobe</sub> lacking almost all region C<sub>IP-II</sub>. Consequently, C<sub>IP-II</sub> is the less conserved part among IP<sub>5</sub> 2-Ks and, in particular, the plant enzyme seems to have evolved to bind Zn that appears to have a structural role. Most of the residues implicated in ligand binding and activity are conserved (Fig. 3A) and are mainly enclosed in motifs I–IV, engaged in both substrates recognition, or in the <sup>192</sup>RxxMHQxLK motif, involved only in the inositol recognition. There is a sixth conserved motif, <sup>310</sup>LDxLDIEGxIHxY not present in fungi, which seems to be essential for maintaining the C<sub>IP-lobe</sub> properly folded. Because fungi have a shorter C<sub>IP-lobe</sub>, this motif and its function are probably not required. Finally, L3 is not conserved among the IP<sub>5</sub> 2-K family, suggesting that L3 is characteristic of each IP<sub>5</sub> 2-K and could be essential in modulation of catalysis.

Two single mutations reported in yeast IP<sub>5</sub> 2-K (N10D and C162Y) yielded inactive enzymes (23). These residues are equivalent to Asn22 and Lys194, located in nucleotide- and inositol-binding regions respectively. To confirm the role of other IP<sub>5</sub>

2-K residues in catalysis, we have performed activity and kinetic analysis of selected mutants (Table 1). R40V activity is half the wild type, probably due to partial loss in phosphate coordination, which would also explain the increase in the K<sub>m</sub> for ATP observed in this mutant. As expected for catalytic residues (see later), mutants K168N and K168A yielded completely inactive enzymes and D407A has only residual activity. The almost complete inactivity of the R130I mutant illustrates the importance of P1 binding and the key role of the N-lobe in the enzymatic activity. Arg130 interacts with L3, and therefore it could also serve to stabilize the inositol binding site conformation. Similarly, Lys170, which coordinates the inositol P5 and P6, is essential for enzymatic function, whereas Gln238, also involved in P6 coordination, is not critical. The K<sub>m</sub> for ATP of N238A mutant is 2- to 3-fold above the wild type. We suspect that the nucleotide-binding pocket is affected by the interaction between L3 and motif II, therefore a change in Asn238 within this motif could have a direct effect in ATP affinity. Asp368, which coordinates 2-OH, yields a completely inactive enzyme when mutated even when assayed at high protein concentration. Asp368 is in turn coordinated to Lys411, a residue that interacts with P3 and P4 of the inositol, and which is also critical for the enzymatic activity. In summary, all the mutagenesis assays support the conclusions deduced from the IP<sub>5</sub> 2-K structures presented here. Finally, it is worth noting that mutant



**Table 1. Activity analysis of IP<sub>5</sub> 2-K mutants**

IP <sub>5</sub> 2-K mutant	Activity (% wild type)	
Native	100.00	
R40V*	56.06	
E85A*	124.24	
R130I	6.06	(9.09)
K168A	nd	(2.27)
K168N	nd	(4.54)
K170S	5.30	(3.78)
N238A*	85.60	
D368A	nd	(nd)
D407A	3.03	(7.57)
K411A	nd	(3.53)

ND: No activity detected. Data in parentheses shows values obtained at 10-fold higher concentration of native.

\*Native ATP Km = 38  $\mu$ M and Vmax = 1080 pmol/min/ $\mu$ g. The Km value increases 2- to 3-fold for the active mutants (R40V, E85A, and N238A), with <20% reduction in Vmax.

E85A (equivalent to Glu91 in PKAs), unexpectedly, has no detrimental effect on activity (see [Fig. S6](#))

**IP<sub>5</sub> 2-K Is a Distant Member of the IPK Family.** A search performed with the DALI server ([http://ekhidna.biocenter.helsinki.fi/dali\\_server/](http://ekhidna.biocenter.helsinki.fi/dali_server/)) shows that IP<sub>3</sub> 3-K isoform C is the third closest structural homologue of IP<sub>5</sub> 2-K (rmsd = 5.4 Å for 153 residues), the C<sub>IP</sub> region not showing any significant structural homologue. We conclude that IP<sub>5</sub> 2-K belongs to the IPKs family, albeit as the most distant member with no structural similarity in the inositide binding region.

IPKs display structural homology with PKs, conserving the N-lobe fold and the nucleotide recognition pattern. However, the C-lobe is divergent and presents some structural relationship with the ATP-grasp fold. All these IPKs features are conserved in IP<sub>5</sub> 2-Ks. Fig. 3A shows how most of the IP<sub>5</sub> 2-K N- and C-lobe  $\beta$ -sheets,  $\alpha$ 4, and, partially, the helices that pack over the sheet, are conserved in IP<sub>3</sub> 3-K, chosen as IPK representative. Particularly, the four motifs I–IV of IP<sub>5</sub> 2-Ks are equivalent to the conserved motifs of IPKs (Fig. 3B and Fig. S7), preserving some of the key residues and functions. Notably, the catalytic residues Lys168 from motif I and Asp407 from motif IV are conserved not only in all IPKs, but also in PKs. In spite of the homology in this region, there are some active conformation hallmarks of PKs (32) present in IPKs, that can not be found in IP<sub>5</sub> 2-Ks (Fig. S6). The conformational changes governing IP<sub>5</sub> 2-K activation might be specific for this family of enzymes.

As mentioned before, the inositide-binding region of IP<sub>5</sub> 2-Ks (C<sub>IP</sub>-lobe) is very different from that of IPKs (IP-lobe). The C<sub>IP</sub>-lobe (190 residues) is much larger than the IP-lobe (30 residues in IPmKs or 60 in IP<sub>3</sub> 3-Ks). The IP<sub>3</sub> 3-Ks IP-lobe is formed by a four-helix insertion that emerges from the C-lobe β1c and β2c strands (Fig. 3, Fig. S7, and Fig. S8A). The equivalent region in IP<sub>5</sub> 2-K is the region C<sub>IP-I</sub> inserted into β7–β8, but presenting a very different fold. The regions C<sub>IP-II</sub> and C<sub>IP-III</sub> have no equivalent in IP<sub>3</sub> 3-K. Curiously, the superposition of IP<sub>5</sub> 2-K and IP<sub>3</sub> 3-K structures reveals that the IP<sub>5</sub> axial and the IP<sub>3</sub> equatorial C-OH acceptor bonds have a similar orientation relative to the nucleotide (Fig. S8). Consequently, the two inositide ring planes are perpendicular to each other. The OH acceptor, is bound to a lysine from motif I in both enzymes (Lys 264 and Lys 168, respectively); however, in IP<sub>5</sub> 2-K it is additionally bound to Asp368 from motif III. In spite of the different ring orientation, the IP<sub>5</sub> P3 falls in a similar spatial disposition to IP<sub>3</sub> P4, both phosphates interacting with a residue from motif IV (Lys411 and Lys419, respectively) (Fig. S8B). No more homology is found in the inositide recognition mode.

Finally, most IPKs have nonconserved N-terminal domains with functions distinct from the kinase activity (Fig. 3B). Only *Schizo*-

*saccharomyces pombe* IP<sub>5</sub> 2-K, has such an N-terminal domain (23). Besides, IPKs have variable insertions in both  $\beta$ -sheets, the most common ones being insert I and insert II. The IP<sub>6</sub> K2 insert I has been involved in interactions with HSP90 (33), and yeast Ipmk insert II has been also implied in protein-protein interactions (34), though other studies discard these interactions (35). The IP<sub>5</sub> 2-K region N-II is located at "insert I" and participates in recognition of the inositol substrate, therefore we postulate that insert I could play also a role in the inositol binding within the IPK family. The IP<sub>5</sub> 2-K region  $\beta$ 11 is located at insert II, but to date, no work implies this  $\beta$ -strand in protein-protein interactions.

## Discussion

We have structural information that shows different snapshots of the IP<sub>5</sub> 2-K catalytic mechanism, that allowed us to identify the residues that coordinate P $\gamma$  and 2-OH before and after phosphoryl transference. These residues are Lys168, Asp407, Ser409, and Asp368, together with some G-loop residues. Lys168 neutralizes the negative charge developed in the transition state, whereas the G-loop, and Asp407 through Mg1, orientates P $\gamma$  for the subsequent nucleophilic attack of the inositol OH. The roles proposed for Lys168 and Asp407 (motifs I and IV) are conserved with the roles of equivalent residues in PKs and IPKs. The IP<sub>5</sub> 2-K specific residues Ser409 (Asp in other species) and Asp368 also seem to participate in the catalytic mechanism. Ser409 coordinates Mg1, possibly assisting in P $\gamma$  orientation and in addition, Ser409 constrains the conformational mobility of L5 and positions Lys411, essential for inositide coordination. Asp 368 recognizes the acceptor 2-OH in axial position of the inositide. It is also linked to Lys168 and Lys411 assisting in the active site conformation (Fig. S3). This residue is topologically equivalent to the Asp262 of the PKA catalytic loop, which selects the correct conformation of the serine substrate. Other IPKs and lipid kinases always bind -OHs in equatorial conformation, and do not have an equivalent acidic residue, supporting the idea that the major role of Asp368 is to select the axial 2-OH substrate. Furthermore, two specific features, the Asp368 2-OH recognition and the whole IP<sub>5</sub> binding pocket, potentially discriminates myo-inositides from other conformers. It remains to be determined whether Asp368 participates in catalysis. After phosphoryl transference, Mg1 reorganizes its coordination, and together with Mg2, bridges nucleotide and inositide. Other potential roles of the metals include assistance in catalysis and stabilization of the transition state or products. The metal roles are conserved with IPKs. Finally, taking into account the short distance covered by the P $\gamma$  to reach its fate (P $\gamma$ -2-OH distance = 3.2 Å; OP $\beta$  - P2 distance = 3.1 Å), we predict that the IP<sub>5</sub> 2-K mechanism involves an “in line” direct transference, as was proposed for IP<sub>3</sub> 3-K and protein kinases.

In conclusion, the structure of an IP<sub>5</sub> 2-K, the only inositol kinase that phosphorylates the axial position of an inositol compound, has been solved. The reported fold serves as a template for the whole IP<sub>5</sub> 2-K family, allowing the identification of the key residues. The determined IP<sub>5</sub> 2-K structure allows its classification as a distant member of the IPK family, with a very different and sophisticated large inositol-binding region (C<sub>IP</sub>-lobe). IP<sub>5</sub> 2-K C<sub>IP</sub>-lobe snugly embraces the phosphates of the *myo*-inositol (1,3,4,5,6)-pentakisphosphate substrate with multiple residues that contribute to substrate orientation. The structure explains the specificity of the catalytic reaction, in which a phosphate group is transferred just to the inositol 2-OH axial position. Finally, a typical insert region of IPKs N-lobe has been implicated in recognition of inositol, allowing us to propose a similar role in other family members. Because IP<sub>5</sub> 2-K represents a key point in the metabolism of highly phosphorylated inositols, detailed knowledge of IP<sub>5</sub> 2-K structure gives very valuable information for cell biology and rationale design of enzyme inhibitors. These inhibitors could facilitate study of the effect of IP<sub>5</sub> 2-K inhibition in mam-

mals, because disruption of this gene yields an embryo lethal phenotype in mice. Furthermore, this structure could be a valuable tool in plant physiology, helping in the design of low phytate crops. Finally, IP<sub>5</sub> 2-K substrate and product, IP<sub>5</sub> and IP<sub>6</sub> have already been reported to confer antitumor effects through PI3K/Akt pathway inhibition. Nevertheless, the most interesting facets of IP<sub>5</sub> 2-K activity manipulation perhaps remain to be investigated, as the product IP<sub>6</sub> participates in multiple signalling processes.

## Methods

**Protein Sample Expression and Purification.** At-IP<sub>5</sub> 2-K was amplified from At5G 42810 cDNA (21) and cloned into the pKLSL<sub>1</sub> vector (36). For the enzyme assays, IP<sub>5</sub> 2-K was cloned into the pOPTG vector. Details of protein expression and purification are found in *SI Methods*.

**Protein Crystallization.** Crystallization procedures were set up as described by us (37). The best crystals were obtained at 18 °C and 22% PEG 3350, 100 mM Bis-Tris pH 5.9, and always in presence of the inositol substrate or product. Details of crystal complexes formation are given in *SI Methods*.

**Data Collection, Phasing, and Model Refinement.** Cryoprotection and heavy atom formation were done as previously reported (37). A standard MAD experiment and native datasets for several complexes were collected (Table S1). Programs used are: MOSFLM and the CCP4 package for data processing and reduction (38, 39); SHELX for heavy atom location (40); autoSHARP for protein phasing (41); Buccaneer for preliminary model building (42); REFMAC for

refinement (43); and O for manual building (44). For model completion we used our best dataset (IP<sub>5</sub> 2-K with products) in combination with the experimental model. A more detailed description is shown in *SI Methods*.

**Enzyme Assays.** Wild-type and mutant enzymes assays and kinetic analysis were performed as shown in *SI Methods*. Products were resolved by reverse-phase HPLC (45). Peak areas for ADP and ATP were integrated after reverse-phase HPLC (40). The effect of EDTA treatment of IP<sub>5</sub> 2K was determined in a coupled assay containing 10 mM-MgCl<sub>2</sub> (46).

**Zinc Quantification and Fluorescence.** Zinc quantification by ICP-OES were performed on native enzyme and fluorescence analysis were performed on native vs. EDTA treated enzyme as shown in *SI Methods*.

**ACKNOWLEDGMENTS.** We thank R. Williams and O. Perisic for critical reading of the manuscript, O. Perisic and J.M. Mancheño for providing us with pOPTG and pKLSL<sub>1</sub> vectors, P. Green for preliminary studies of enzyme activity, Chris Hamilton for assistance with assays, and Graham Chilvers for performing ICP-OES analysis. We thank the European Synchrotron Radiation Facility for providing time and assistance for data collection. B.G. was supported by Ramon y Cajal Fellowship RYC-2006-002701. J.I.B. was supported by Formación de Profesorado Universitario (FPU) Fellowship AP2008-00916 from the Ministerio de Educación. C.A.B. and P. Green were supported by Biotechnology and Biological Sciences Research Council of the UK Grant BB/C514090/1. The work was supported by Comunidad de Madrid-CSIC Grants CCG07-CSIC/GEN-2232 and CCG08-CSIC/GEN-3490, Ministerio de Ciencia e Innovación Grant BFU2008-02897/BMC, and Ministerio de Educación y Ciencia Grant RYC-2006-002701.

1. Michell RH (2008) Inositol derivatives: Evolution and functions. *Nat Rev Mol Cell Biol* 9:151–161.
2. Irvine RF, Schell MJ (2001) Back in the water: The return of the inositol phosphates. *Nat Rev Mol Cell Biol* 2:327–338.
3. York JD, Odom AR, Murphy R, Ives EB, Wente SR (1999) A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science* 285:96–100.
4. Hanakahi LA, Bartlett-Jones M, Chappell C, Pappin D, West SC (2000) Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. *Cell* 102:721–729.
5. Murphy AM, Otto B, Brearley CA, Carr JP, Hanke DE (2008) A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *Plant J* 56:638–652.
6. Agarwal R, Mumtaz H, Ali N (2009) Role of inositol polyphosphates in programmed cell death. *Mol Cell Biochem* 328:155–165.
7. Shen X, Xiao H, Ranallo R, Wu WH, Wu C (2003) Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* 299:112–114.
8. Steger DJ, Haswell ES, Miller AL, Wente SR, O'Shea EK (2003) Regulation of chromatin remodeling by inositol polyphosphates. *Science* 299:114–116.
9. Saiardi A, Caffrey JJ, Snyder SH, Shears SB (2000) The inositol hexakisphosphate kinase family. Catalytic flexibility and function in yeast vacuole biogenesis. *J Biol Chem* 275:24686–24692.
10. Mulugu S, et al. (2007) A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science* 316:106–109.
11. Shukla VK, et al. (2009) Precise genome modification in the crop species Zea mays using zinc-finger nucleases. *Nature* 459:437–441.
12. Raboy V (2003) myo-Inositol-1,2,3,4,5,6-hexakisphosphate. *Phytochemistry* 64:1033–1043.
13. Raboy V (2001) Seeds for a better future: "Low phytate" grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci* 6:458–462.
14. Raboy V (2007) The ABCs of low-phytate crops. *Nat Biotechnol* 25:874–875.
15. Shamsuddin AM (2008) Demineralizing phytate. *Nat Biotechnol* 26:496–497.
16. Bozsik A, Kokeny S, Olah E (2007) Molecular mechanisms for the antitumor activity of inositol hexakisphosphate (IP<sub>6</sub>). *Cancer Genomics Proteomics* 4:43–52.
17. Graf E, Empson KL, Eaton JW (1987) Phytic acid. A natural antioxidant. *J Biol Chem* 262:11647–11650.
18. Maffucci T, et al. (2005) Inhibition of the phosphatidylinositol 3-kinase/Akt pathway by inositol pentakisphosphate results in antiangiogenic and antitumor effects. *Cancer Res* 65:8339–8349.
19. Stephens LR, et al. (1991) myo-Inositol pentakisphosphates. Structure, biological occurrence and phosphorylation to myo-Inositol hexakisphosphate. *Biochem J* 275:485–499.
20. Phillippy BQ, Ullah AH, Ehrlich KC (1994) Purification and some properties of inositol 1,3,4,5,6-pentakisphosphate 2-kinase from immature soybean seeds. *J Biol Chem* 269:28393–28399.
21. Sweetman D, Johnson S, Caddick SE, Hanke DE, Brearley CA (2006) Characterization of an Arabidopsis inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). *Biochem J* 394:95–103.
22. Verbsky JW, Wilson MP, Kisseleva MV, Majerus PW, Wente SR (2002) The synthesis of inositol hexakisphosphate. Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *J Biol Chem* 277:31857–31862.
23. Ives EB, Nichols J, Wente SR, York JD (2000) Biochemical and functional characterization of inositol 1,3,4,5, 6-pentakisphosphate 2-kinases. *J Biol Chem* 275:36575–36583.
24. Sun Y, et al. (2007) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize: molecular and biochemical characterization. *Plant Physiol* 144:1278–1291.
25. Verbsky J, Lavine K, Majerus PW (2005) Disruption of the mouse inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression. *Proc Natl Acad Sci USA* 102:8448–8453.
26. Sarmah B, Latimer AJ, Appel B, Wente SR (2005) Inositol polyphosphates regulate zebrafish left-right asymmetry. *Dev Cell* 9:133–145.
27. Gonzalez B, et al. (2004) Structure of a human inositol 1,4,5-trisphosphate 3-kinase: substrate binding reveals why it is not a phosphoinositide 3-kinase. *Mol Cell* 15:689–701.
28. Miller GJ, Hurley JH (2004) Crystal structure of the catalytic core of inositol 1,4,5-trisphosphate 3-kinase. *Mol Cell* 15:703–711.
29. Miller GJ, Wilson MP, Majerus PW, Hurley JH (2005) Specificity determinants in inositol polyphosphate synthesis: crystal structure of inositol 1,3,4-trisphosphate 5/6-kinase. *Mol Cell* 18:201–212.
30. Cheek S, Ginalski K, Zhang H, Grishin NV (2005) A comprehensive update of the sequence and structure classification of kinases. *BMC Struct Biol* 5:6.
31. Stevenson-Paulik J, Bastidas RJ, Chiou ST, Frye RA, York JD (2005) Generation of phytate-free seeds in *Arabidopsis* through disruption of inositol polyphosphate kinases. *Proc Natl Acad Sci USA* 102:12612–12617.
32. Nolen B, Taylor S, Ghosh G (2004) Regulation of protein kinases; controlling activity through activation segment conformation. *Mol Cell* 15:661–675.
33. Chakraborty A, et al. (2008) HSP90 regulates cell survival via inositol hexakisphosphate kinase-2. *Proc Natl Acad Sci USA* 105:1134–1139.
34. Messenguy F, Dubois E, Boonchird C (1991) Determination of the DNA-binding sequences of ARGR proteins to arginine anabolic and catabolic promoters. *Mol Cell Biol* 11:2852–2863.
35. Caddick SE, Harrison CJ, Stavridou I, Johnson S, Brearley CA (2007) A lysine accumulation phenotype of ScpK2Delta mutant yeast is rescued by Solanum tuberosum inositol phosphate multikinase. *Biochem J* 403:381–389.
36. CSIC (2009) PCT Patent Application WO/2009/121994.
37. Baños-Sanz JI, Villate M, Sanz-Aparicio J, Brearley CA, Gonzalez B (2009) Crystallization and preliminary x-ray diffraction analysis of inositol 1,3,4,5,6-pentakisphosphate kinase from *Arabidopsis thaliana*. *Acta Crystallogr F* 66:102–106.
38. Leslie AGW (1992) Joint CCP4 + ESF-EAMCB Newsletter Protein Crystallogr. (Daresbury Laboratory, Warrington, UK), 26.
39. CCP4 (Collaborative Computational Project Number 4) (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
40. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
41. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
42. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D* 62:1002–1011.
43. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
44. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
45. Caddick SE, et al. (2008) A Solanum tuberosum inositol phosphate kinase (StIPK1) displaying inositol phosphate-inositol phosphate and inositol phosphate-ADP phosphotransferase activities. *FEBS Lett* 582:1731–1737.
46. Rowan AS, et al. (2009) Nucleoside triphosphate mimicry: A sugar triazoly nucleoside as an ATP-competitive inhibitor of *B. anthracis* pantothenate kinase. *Org Biomol Chem* 7:4029–4036.

# Supporting Information

González et al. 10.1073/pnas.0912979107

## SI Text

**SI Methods. Protein sample expression and purification.** IP<sub>5</sub> 2-K from *Arabidopsis thaliana* was amplified from At5G 42810 cDNA and cloned into the pKLSL<sub>1</sub> vector (1). The fused protein was expressed in *Escherichia coli* Rosetta pLysS strain, growing the cells at 37 °C until OD<sub>600</sub> = 1 and induced by 0.3 mM of IPTG for 15 hours at 16 °C. Clarified cell lysate in buffer A (20 mM Tris/HCl pH 8, 150 mM NaCl and 2 mM DTT) was applied to a Sepharose CL-6B column and washed with buffer A. The protein was eluted by buffer A + 200 mM lactose. The sample was diluted, loaded onto a heparin column, washed by buffer B (20 mM Tris/HCl pH 8, 50 mM NaCl and 2 mM DTT) and eluted with a salt gradient. The fused protein was cleaved with Tobacco Etch Virus (TEV) protease at 4 °C over night. The sample was further purified by gel filtration on a 16/60 Superdex 200 column in buffer C (20 mM Tris/HCl pH 8, 200 mM NaCl and 2 mM DTT). The residual fusion protein was removed using a small Sepharose CL-6B column. The final pure protein was concentrated to 10 mg/mL and stored at -80 °C.

For the enzyme assays, IP<sub>5</sub> 2-K was cloned into the pOPTG vector. GST-mutants were prepared by Quick Change site directed mutagenesis (Stratagene). Expressions were carried out as before. The lysates in buffer B were loaded on an heparin column, washed by buffer B, and eluted with a salt gradient. The sample was loaded on a Glutathione Sepharose HP column, washed with buffer D (50 mM Tris/HCl pH 7.5, 200 mM NaCl and 2 mM DTT) and cleaved on-column with TEV protease overnight at 4 °C. The protein was eluted and frozen at -80 °C. This procedure did not yield any amount of the Zn site mutants, although fusion proteins were detected after the heparin column. To show the TEV effect over time, the heparin captured fusion proteins were incubated with Glutathione Sepharose 4B beads equilibrated in buffer D and, after washing, they were left rolling o/n in presence of TEV protease. An SDS-PAGE analysis was done.

All constructs produced have been sequenced. The protein used for our studies has some discrepancies with the AT5G42810 sequence annotated in the database. None of these residues are relevant and the native enzymes produced are fully active.

**Protein crystallization.** The best crystals were obtained in drops of 2  $\mu$ L of protein sample and 2  $\mu$ L of precipitant, using the sitting drop procedure, at 18 °C and 22% PEG 3350, 100 mM Bis-Tris pH 5.9 (2 mM MgCl<sub>2</sub> was also added to saturate metal positions), always in presence of inositol. The protein was crystallized with the substrate, 2 mM IP<sub>5</sub> (Alexis), the product, 2 mM IP<sub>6</sub> (Sigma), substrate and analogue, 2 mM IP<sub>5</sub> and 2 mM AMPPNP (Sigma), and both products, 2 mM IP<sub>6</sub> and 2 mM ADP (Sigma). Crystals appeared as rod clusters in 3–7 days and kept growing for two weeks.

**Data collection, phasing and model refinement.** For data collection individual crystals were broken from clusters and cryoprotected by changing gradually the drop solution to 35% PEG 3350 and 100 mM Bis-Tris pH 5.9. For the heavy atom derivative formation, the cryoprotectant solution plus 100 mM of lead acetate was added to crystals drops obtained with both products (ADP and IP<sub>6</sub>). The crystals were incubated in this solution for 25 min. Data collections were carried out at 100 K in European Synchrotron Radiation Facility (ESRF) synchrotron beamlines. In a standard multiwavelength anomalous dispersion experiment and native datasets for several complexes were collected (Table S1).

Data processing and reduction was performed using MOSFLM (2) and the CCP4 package (3). Four lead peaks were located by SHELX (4), refined and used for protein phasing with autoSHARP (5). Solvent flattening was carried out with solvent content of 43.2%, suggesting the presence of two molecules in the crystal asymmetric unit. A preliminary incomplete model was built using Buccaneer program (6). For model completion we used our best dataset (IP<sub>5</sub> 2-K with products) in combination with the experimental model. Refinement was carried out using REFMAC (7) alternated with cycles of manual building using O (8). The structure of other complexes was obtained by Fourier difference map analysis. Final statistics and figures of merit are given in Table S1. PROCHECK reveals that 91% of the protein structure residues are in the most favorable regions of the Ramachandran plot. Small regions of molecule A (49–58, 156–157, 336–341, 438–452) and molecule B (54–56, 156–157, 336–341, 438–452) are not ordered in the structure. The figures were generated with PyMOL.

**Enzyme assays.** Wild-type and mutant enzymes were assayed in buffer E (10 mM HEPES pH 7.5, 6 mM MgCl<sub>2</sub>, 2 mM EDTA and 2 mM DTT) containing 0.2 mM ATP, 0.1 mM Ins(1,3,4,5,6)P<sub>5</sub>, 3.7 kBq of [ $\gamma$ -<sup>33</sup>P]ATP and 0.08  $\mu$ g protein in 10  $\mu$ L volume. Reactions (5 min at 37 °C) were limited to approx 10% conversion of ATP. Analyses were performed in triplicate and the whole experiment repeated twice. For mutants with low activity, the assay was performed additionally at 10 $\times$  higher protein concentration. Products were resolved by reverse-phase HPLC (4). The enzymes did not release inorganic phosphate, ie. did not show phosphatase activity.

For kinetic analysis, InsP<sub>5</sub> 2-kinase activity was studied as InsP<sub>5</sub>-dependent production of ADP with 0.1 mM IP<sub>5</sub> and ATP in the approximate range 0.01–0.2 mM. Assays were performed as above; peak areas for ADP and ATP were integrated after reverse-phase HPLC (4). In the absence of InsP<sub>5</sub>, no ADP production was observed.

The effect of EDTA-treatment of Ipk1 protein was determined in a coupled assay after Rowan et al. (9). The assay (0.7 mL) contained 450 ng protein, 116  $\mu$ M-ATP, 11.6  $\mu$ M-Ins(1,3,4,5,6)P<sub>5</sub> and 10 mM MgCl<sub>2</sub>. Protein was preincubated for 30 min at 4 °C with 2 mM DTT, 50 mM NaCl, 50 mM Tris/HCl, pH 7.3 (control) or 2 mM DTT, 50 mM EDTA (disodium salt), 50 mM-Tris/HCl, pH 7.3 (EDTA-treated) prior to dilution (90-fold) in the assay. Reactions were monitored at 340 nm in a Cary 50 UV/Vis spectrophotometer. The experiment was repeated a further two times with similar results.

**Fluorescence.** Fluorescence emission spectra were recorded on a Cary Eclipse fluorescence spectrometer. IP<sub>5</sub> 2-K protein (28 ng) was added from 1.13 mg/mL stock to 1 mL of 2 mM-DTT, 50 mM-NaCl, 50 mM-Tris/HCl, pH 7.3, in a 3 mm  $\times$  10 mm quartz glass cuvette. The cuvette contents were maintained at 25 °C. The excitation wavelength was 295 nm and the band pass for excitation and emission was set to 5 nm. The emission scan rate was 30 nm/min and the photomultiplier tube (PMT) voltage was set to 800 V. EDTA was added in 10  $\mu$ L aliquots from a solution of 50 mM EDTA (disodium salt) made up in 2 mM DTT, 50 mM Tris/HCl, pH 7.3. The pH of this solution was adjusted to pH 7.3 with NaOH. All spectra are corrected for dilution of protein on addition of EDTA. Fluorescence changes were complete within 60 s. Second spectra run 5 min after the first

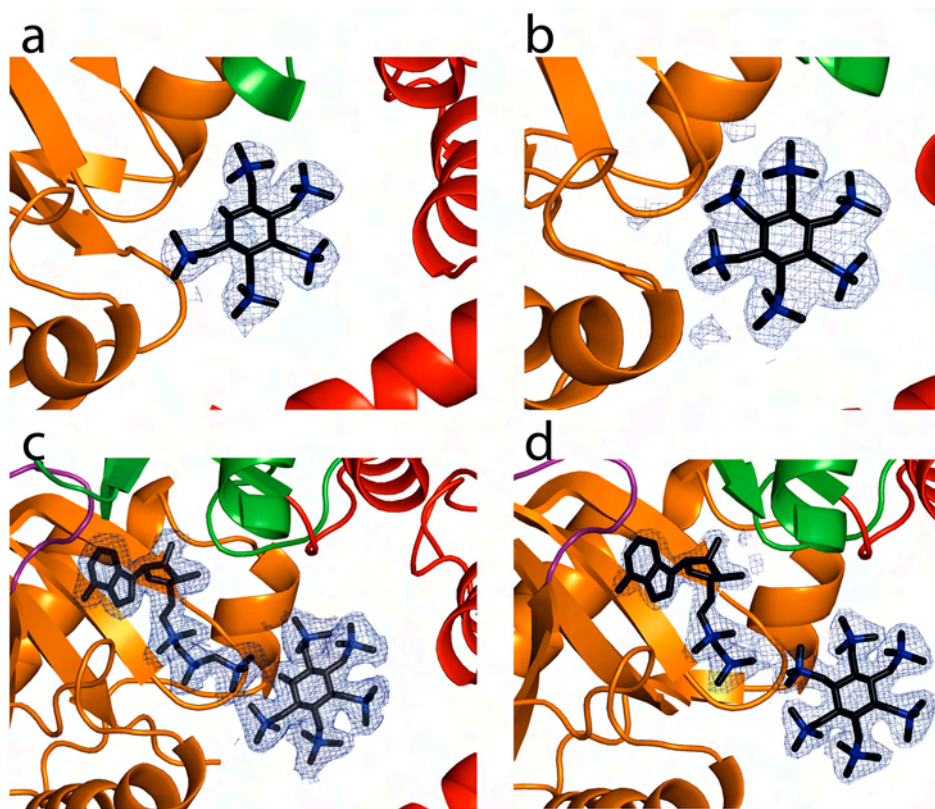


spectra were superimposable on the first spectra obtained after each addition.

**Inductively Coupled Plasma-Optical Emission Spectroscopy** Metal analysis of IP<sub>5</sub> 2-K was performed by optical emission spectroscopy on a Varian Vista Pro ICP-OES. Protein (0.2 mL of a 1.13 mg/mL solution) was diluted to 2 mL with 50 mM Tris/HCl, pH 7.3. The sample was applied to a centrifugal microconcentrator [Vivaspin™2 (Sigma-Aldrich) centrifugal concentrators MWCO 10,000] and centrifuged for 10 minutes at 7500 × g at a temperature of 20 °C. This process was repeated twice and a further three times with 50 mM Tris/HCl, pH 7.3. The concentrated protein sample was resuspended to 0.5 mL with 5 mM Tris/HCl, pH 7.3 and further diluted 20× in the same

buffer. Diluted protein was subjected to ICP-OES on a machine calibrated with 0–4.0  $\mu\text{M}$  standards of Co, Cu, Ni, and Zn. Individual Vivaspin<sup>™</sup> eluates and the 5 mM Tris/HCl, pH 7.3 buffer were analyzed also. The Zn content of the final processed protein sample was 0.57  $\mu\text{M}$  and that of the buffer 0.12  $\mu\text{M}$ , giving a zinc content of 9.0  $\mu\text{M}$  attributable to protein prior to final dilution. The protein content of the nondiluted, washed protein sample was determined by Bradford assay and estimated to be  $9.4 \pm 0.6 \mu\text{M}$ . The Co, Cu, and Ni content of the final protein sample was 16, 14, and 115 nM, respectively, and 13, not detectable and 124 nM, respectively, for the buffer. This data indicates a stoichiometry of one zinc atom per protein molecule and is consistent with the structural analysis.

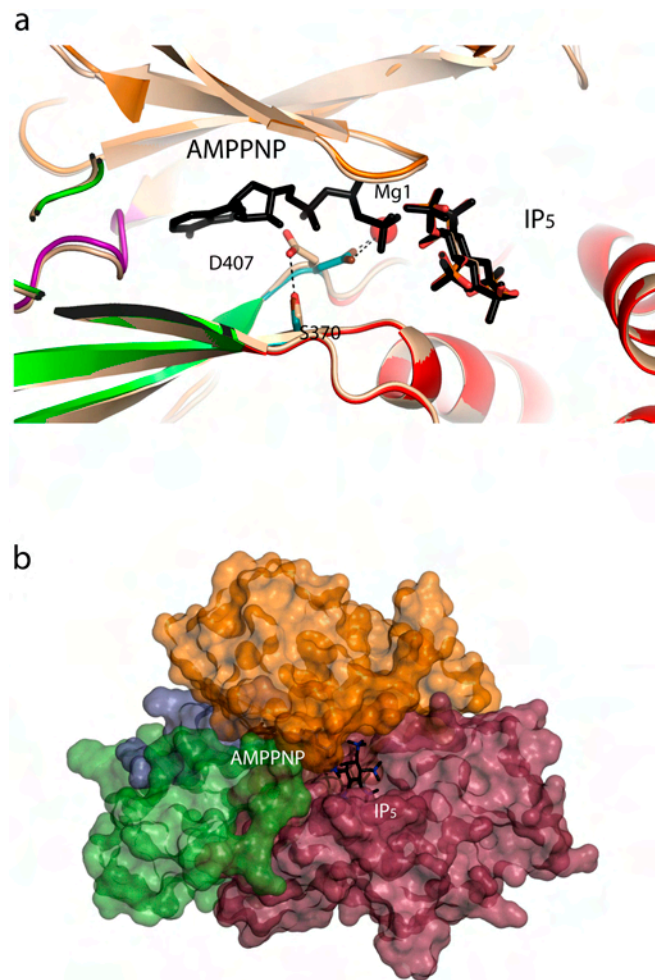
1. CSIC (2009) *PCT Patent Application W/2009/121994*.
2. Leslie AGW (1992) *Joint CCP4 + ESF-EAMCB Newslett Protein Crystallogr* (Daresbury Laboratory, Warrington, UK), vol. 26.
3. CCP4 (Collaborative Computational Project Number 4) (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
4. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
5. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
6. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D* 62:1002–1011.
7. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
8. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
9. Rowan AS, et al. (2009) Nucleoside triphosphate mimicry: A sugar triazolyl nucleoside as an ATP-competitive inhibitor of *B. anthracis* pantothenate kinase. *Org Biomol Chem* 7:4029–4036.



**Fig. S1.** Electron density map for ligands. 2fo- $\sigma$  electron density map for the ligands in the different complexes solved (*a*): IP<sub>5</sub> 2-K/IP<sub>5</sub>; *b*): IP<sub>5</sub> 2-K/IP<sub>6</sub>; *c*): IP<sub>5</sub> 2-K/AMPPNP/IP<sub>5</sub>, and *d*): IP<sub>5</sub> 2-K/ADP/IP<sub>6</sub>) contoured at 2 $\sigma$  level.

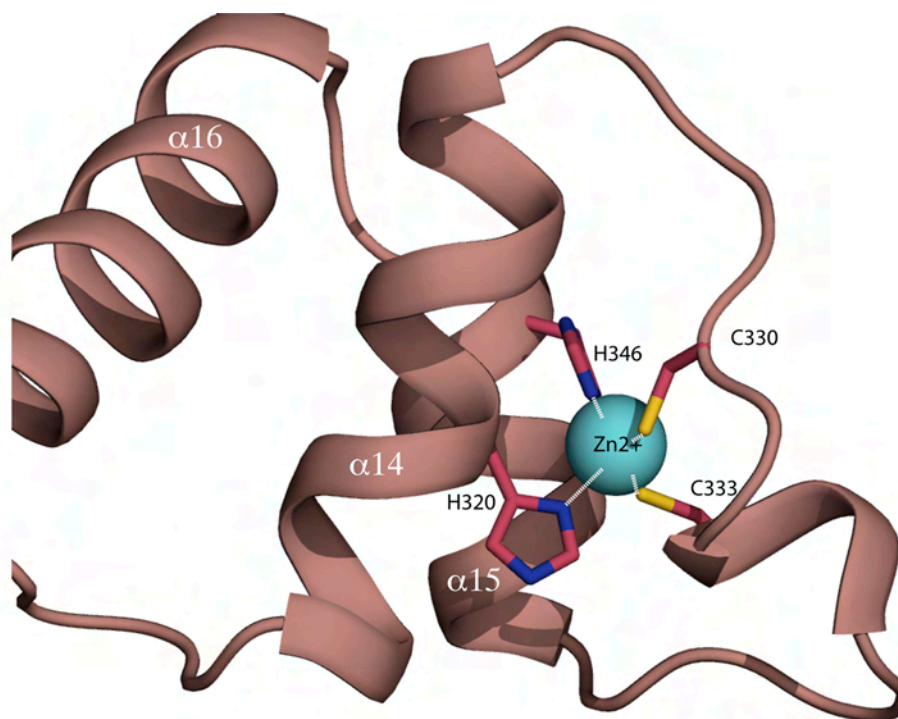


**Fig. S3.** IP<sub>5</sub> 2-K ligands interactions. Scheme of the interactions between IP<sub>5</sub> 2-K and the substrates IP<sub>5</sub> and AMPPNP, prepared with ChemDraw.



**Fig. 54.** Binary and ternary complexes. (a) Superposition of the active site of IP<sub>5</sub> 2-K/IP<sub>5</sub> and IP<sub>5</sub> 2-K/AMPPNP/IP<sub>5</sub> complexes, showing the Asp407 side chain variation, shown in cream and cyan, respectively. Asp407 only coordinates Mg in the ternary complex. (b) IP<sub>5</sub> 2-K surface, showing how the inositide seems to limit the available room for the nucleotide entrance.

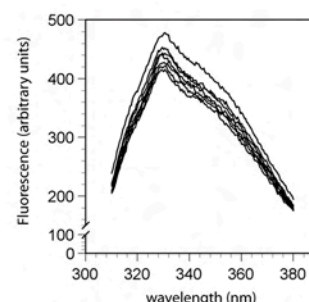
a



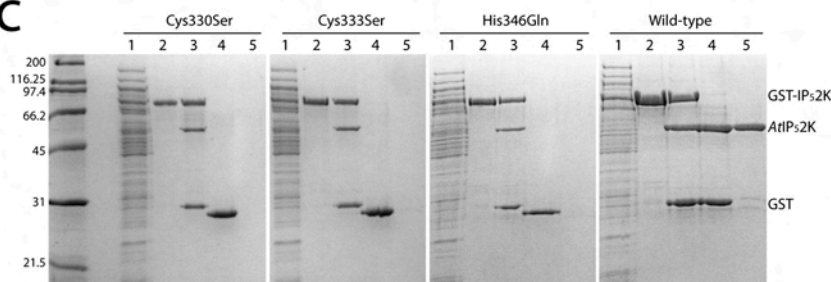
**b**

$\alpha 13$   $\alpha 14$   
 $\alpha 13$   $\alpha 14$   
IP52K\_At 299 GVLDRLLLEIQKLDKLDIEGAIHCYYDIINQCPICKE  
IP52K\_Md 311 GVLDRLLLEAQKLDNDIEGAIHAYDYIISEPMVCGE  
IP52K\_Os 291 GVLGKLLATKLDLDHIEGAIHLYYNIISQPLCVCKN  
IP52K\_Sb 293 GVLGKLLTTQKLDLDHIEGAIHLYYNIISQPLCVCKN  
IP52K\_Zm 293 EVLGNLLATKLDLDHIEGVHLYYNIISQPLCVCKN  
  
 $\alpha 15$   $\alpha 16$   
 $\alpha 15$   $\alpha 16$   
IP52K\_At 336 G...RPLEAELSLALPLDSESKIVEYLAATAKDC  
IP52K\_Md 348 LGKEKESQKYLSTLSMSLDSEKIMVDFLISATAKDC  
IP52K\_Os 328 ITDELLRKYSTLSLSPLDKSEKIVRDFLISATAKDC  
IP52K\_Sb 330 LTDVELLRKYTLHSLPLDKSKIVRDFLISATAKDC  
IP52K\_Zm 330 LTDVELLRKYTLHSLPLDKSKIVRDFLISATAKDC

d



C



e

	Mean (n=5)	sd
Control	8.93	0.73
EDTA-treated	10.33	1.83

**Fig. S5.** The zinc binding motif of *A. thaliana* IP<sub>5</sub> 2-K. (a) Zinc binding motif of *A. thaliana* IP<sub>5</sub> 2-K as found in the product ternary complex, showing the four Zn ligands (His320, Cys330, Cys333, and His346) in sticks representation. (b) Alignment of IP<sub>5</sub> 2-Ks from *Arabidopsis thaliana*, *Malus domestica*, *Oryza sativa*, *Sorghum bicolor*, and *Zea mays* in the C<sub>IP-II</sub> region, showing the Zn ligands conservation (c) SDS-PAGE analysis of C330S, C333S, and H346N IP<sub>5</sub> 2-K mutants and wild type, monitoring the heparin bound material (1), GST-4B bead-bound protein (2), beads + supernatant after 2 h (3) and after o/n (4) TEV incubation, and supernatant after o/n TEV incubation (5). There is no soluble protein in mutants after complete cleavage. Mutant fractions in lanes 1–5 are 3–fold more concentrated than native to make visible the IP<sub>5</sub> 2-K band after 2 h TEV incubation (d) Fluorescence emission spectra of IP<sub>5</sub> 2-K protein and of protein titrated with EDTA (0.5–2.5 mM) in 0.5 mM increments. Successive additions caused incremental decreases in fluorescence. A subsequent addition of 5 mM-MgCl<sub>2</sub> did not alter the fluorescence properties of IP<sub>5</sub> 2-K. Subsequent addition of 1.25 mM ZnCl<sub>2</sub> caused IP<sub>5</sub> 2-K to precipitate as evidenced by scattering of excitation light into the emission spectra and loss of fluorescence after clarification of the sample by centrifugation. (e) Estimation of rates of the reaction (nmol/ min /μg at 25 °C) for the control and EDTA-treated enzyme.

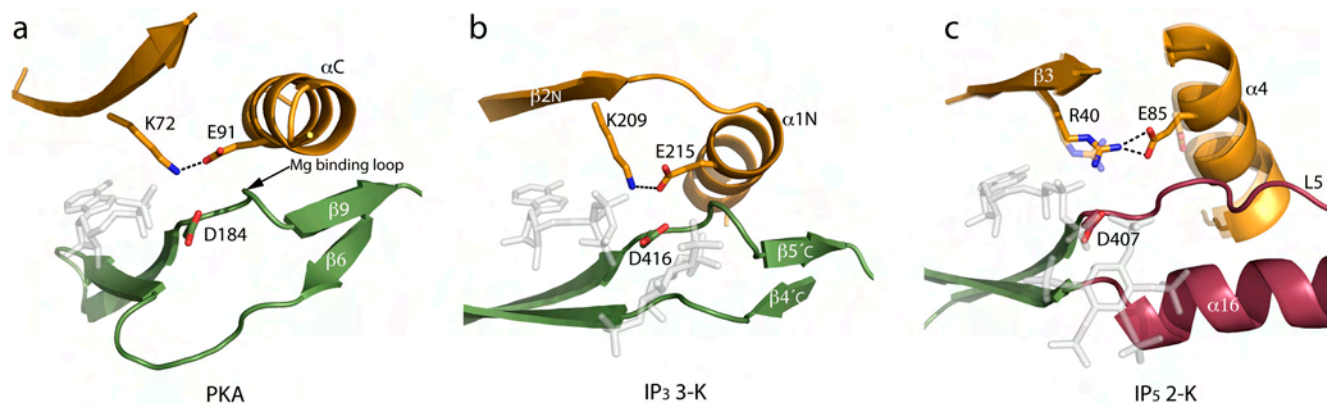
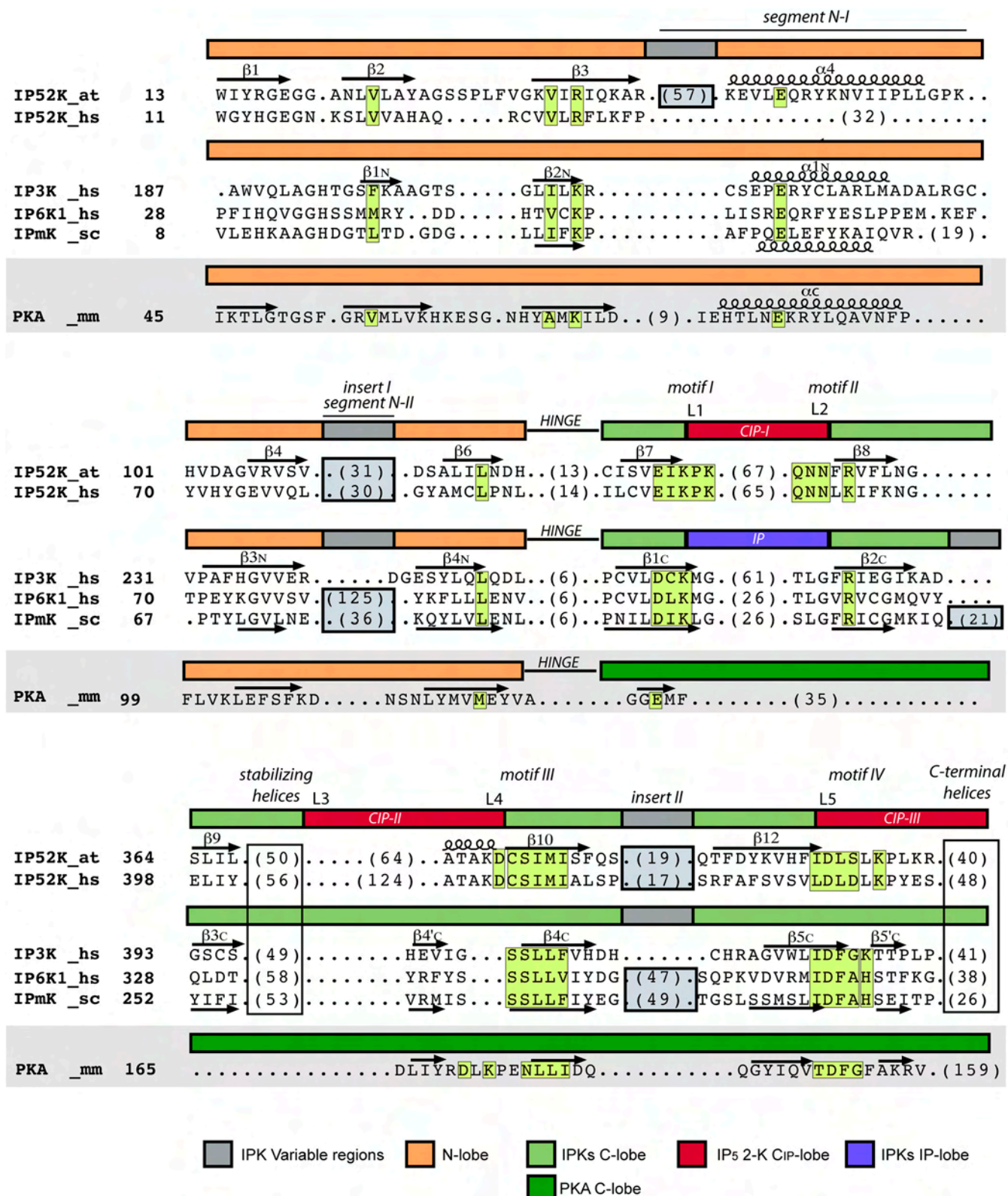


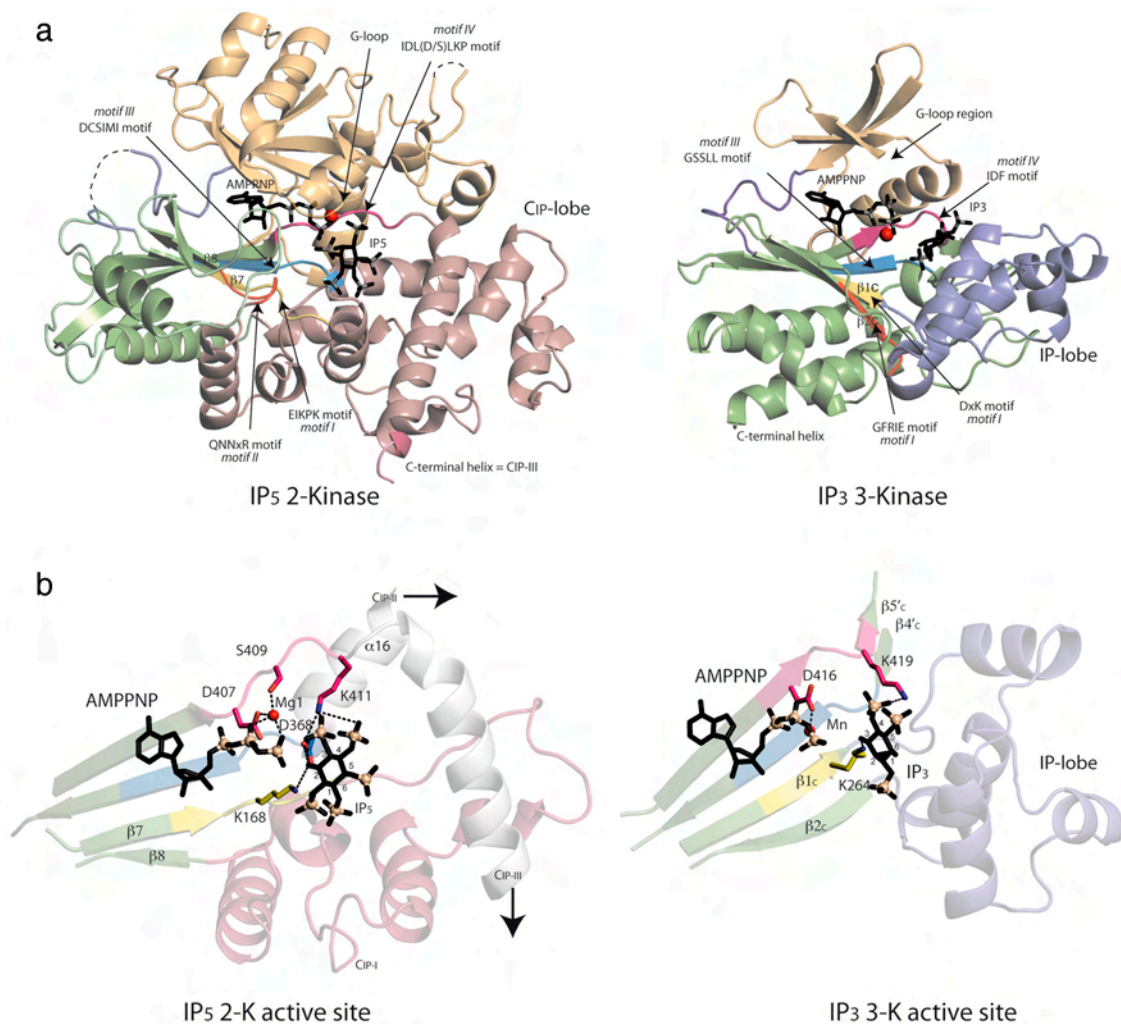
Fig. S6.

Active conformation hallmarks of PKs are absent in IP<sub>5</sub> 2-Ks. (a) PKs in active conformation display a conserved salt bridge in the N-lobe (Lys72-Glu91 in PKA), associated with a proper orientation of  $\alpha$ C. The basic residue is involved nucleotide phosphates binding. In addition, the presence of a small  $\beta$ -sheet in PKs ( $\beta$ 6 and  $\beta$ 9 in PKA) forces a proper conformation of the magnesium binding loop. (b) IP<sub>3</sub> 3-K, as a representative of the IPK family, conserves both hallmarks. (c) IP<sub>5</sub> 2-K does not follow this pattern. First, the equivalent of the salt bridge (Arg40-Glu85) is not always formed in the crystals, and its presence does not produce a significant change in protein conformation. In addition, E85A mutation does not reduce kinase activity (Table 1), supporting a noncritical role for this residue. Interestingly, all other reported IPKs and PKs present a lysine instead an arginine in this position, which may account for this difference. Second, the presence of  $\alpha$ 16 seems to preclude the small beta sheet formation, unless a large conformational change is produced. Nevertheless, the interaction formed between  $\alpha$ 16 and the equivalent of  $\beta$ 5'C region (L5) might have a similar function.





**Fig. S7.** Alignment of IPK family. Structural alignment of IP<sub>5</sub> 2-Ks vs. the IPKs (schematized in Fig. 3B). Mouse PKA as representative of PKs has also been included (pdb code: 1 atp). A representative for each of the three IPK subfamilies has been chosen [human IP<sub>3</sub> 3-K (pdb code:1 w2c)], human IP<sub>6</sub> Kinase (IP<sub>6</sub> K), and yeast InsP<sub>3</sub> multikinase (IPmK) (pdb code: 2iew). Sequences from human IP<sub>5</sub> 2-Ks and IP<sub>6</sub> Ks are aligned based on their secondary structure predictions. In general, the N-lobe and some C-lobe elements are conserved within the IPK family. However, regions involved in inositol binding (CIP-lobe and IP-lobe) are not equivalent. PKA conserves the N-lobe, and a few C-lobe elements involved in catalysis and nucleotide binding. Relevant residues and equivalent motifs are highlighted.



**Fig. 58.** IP<sub>5</sub> 2-K vs. IP<sub>3</sub> 3-K. (a). *A. thaliana* IP<sub>5</sub> 2-K (Left) and human IP<sub>3</sub> 3-K (Right) structures. The four conserved motifs in IP<sub>5</sub> 2-Ks with participation in catalysis and their equivalent motifs in IP<sub>3</sub> 3-Ks are shown in different colors: hot pink (motif IV), sky blue (motif III), yellow (motif I), and red (motif II). The position of the nonequivalent C-terminal helices are also shown. (b) Zoom of the IP<sub>5</sub> 2-K (Left) and IP<sub>3</sub> 3-K (Right) active sites. Residues important for enzymatic activity are shown as sticks. IP<sub>5</sub> 2-K conserves the catalytic Asp407 and Lys168, and presents two new residues, involved in metal coordination (Ser409) and 2-OH coordination (Asp368). A basic residue from L5 recognizes an inoside phosphate contiguous to the acceptor -OH in both enzymes.



**Table S1. Data collection, structure determination, and refinement**

Dataset	PbAc- $\lambda$ 1 peak	PbAc- $\lambda$ 2 edge	PbAc- $\lambda$ 3 remote	AMP PNP-IP <sub>5</sub>	ADP-IP <sub>6</sub>	IP <sub>6</sub>	IP <sub>5</sub>
Data collection							
Cell (Å)		58.06		57.42	58.12	57.67	58.76
		110.97		113.43	113.59	112.50	112.74
		138.69		140.47	142.48	139.72	144.45
Wavelength (Å)	0.94780	0.95030	0.94420	0.93950	0.93340	0.93300	0.95380
Resolution range (Å)	138.67–3.20	140.03–3.20	138.70–3.20	88.25–2.20	71.24–2.20	112.50–3.10	88.74–2.90
Rmerge (%)	15.4 (41.5)	19.0 (55.7)	16.5 (45.3)	9.9 (46.3)	13.5 (53.2)	11.6 (43.6)	13.0 (44.6)
Unique reflection	15450	15656	15489	47247	48686	17187	22033
Redundancy (%)	6.7 (6.7)	6.7 (6.8)	6.7 (6.8)	7.6 (7.6)	7.1 (5.7)	7.6 (7.9)	6.9 (7.0)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.9 (99.4)	99.6 (99.6)	99.8 (99.8)	100.0 (100.0)	100.0 (100.0)
Mn (I/sd)	10.5 (4.8)	9.6 (3.8)	10.6 (4.7)	16.3 (4.4)	10.4 (3.0)	13.7 (4.9)	13.7 (3.8)
I/ $\sigma$ (I)	4.3 (1.7)	3.6 (1.2)	4.0 (1.5)	6.1 (1.6)	3.8 (1.4)	5.8 (1.8)	4.2 (1.7)
Phasing statistics							
Pb sites found	4						
FOM (acent./cent.)	0.42711/0.24412						
Rcullis							
ANO (acentric)	0.884	0.938	0.889				
ISO (acent./centric)	-	0.769/0.794	0.748/0.797				
Phasing power							
ANO (acentric)	0.856	0.568	0.843				
ISO (acent./centric)	-	0.421/0.370	0.019/0.016				
Refinement statistics							
Resolution range (Å)	86.71–3.20			140.03–2.20	88.73–2.20	112.50–3.10	88.74–2.90
Rfactor* (%)	21.0 (26.5)			22.4 (23.9)	22.8 (26.9)	24.1 (30.6)	22.4 (29.7)
R-free† (%)	26.8 (30.9)			27.7 (29.0)	26.9 (31.0)	29.8 (35.0)	29.0 (38.0)
All atoms (non-H)	6851			6991	7015	6751	6668
Protein atoms	6719			6867	6694	6671	6602
Ligand atoms	124			124	124	72	64
Ions	6			4	6	2	2
Waters	0			259	193	0	0
Rms deviations							
Bonds (Å)	0.011			0.006	0.007	0.008	0.010
Angles (°)	1.462			1.109	1.149	1.190	1.371
Mean B /Wilson (Å <sup>2</sup> )	32.8/53.8			29.0/29.1	34.6/36.7	54.7/67.1	41.1/57.0
Protein	32.4			29.3	35.2	55.8	41.3
Ligands	13.2			18.2	21.3	29.1	21.7
Ramachandran data							
Most favored (%)	86.1			91.2	91.0	89.9	91.4

Crystals belong to space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>. Data for the outermost shell are given in parentheses

\*Rmerge =  $\sum_{hkl} \sum_i |I_i(hkl) - \langle I(h, k, l) \rangle| / \sum_{hkl} \sum_i I_i(h, k, l)$

†Rfactor =  $\sum |F_{obs} - F_{calc}| / \sum F_{obs}$  (R-free is the R value for a 5% of reflections randomly selected and that were not used during refinement).

Jose Ignacio Baños-Sanz,<sup>a</sup>  
Julia Sanz-Aparicio,<sup>a</sup> Charles A.  
Brearley<sup>b</sup> and Beatriz  
González<sup>a\*</sup>

<sup>a</sup>Grupo de Cristalografía Macromolecular y  
Biología Estructural, Instituto de Química-Física  
'Rocasolano', CSIC, Serrano 119,  
28006 Madrid, Spain, and <sup>b</sup>School of  
Biological Sciences, University of East Anglia,  
Norwich NR4 7TJ, England

Correspondence e-mail: xbeatriz@iqfr.csic.es

Received 16 March 2012

Accepted 18 April 2012

## Expression, purification, crystallization and preliminary X-ray diffraction analysis of the apo form of InsP<sub>5</sub> 2-K from *Arabidopsis thaliana*

Inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IP<sub>5</sub> 2-K) is a key enzyme that catalyzes the synthesis of phytic acid (IP<sub>6</sub>) from inositol 1,3,4,5,6-pentakisphosphate (IP<sub>5</sub>) and ATP. The first structure of IP<sub>5</sub> 2-K, that from *Arabidopsis thaliana*, has been solved previously; it only crystallized in the presence of inositol, either the substrate IP<sub>5</sub> or the product IP<sub>6</sub>, and failed to crystallize in its free state (without inositol). Based on structural analysis, a point mutation of IP<sub>5</sub> 2-K (W129A) has been produced in order to overcome this limitation and obtain information about protein conformational changes upon substrate binding. Here, the production and crystallization of W129A IP<sub>5</sub> 2-K in its free state and with bound nucleotide is described. These crystals differed from the native crystals and belonged to the orthorhombic space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub>, with unit-cell parameters *a* = 66.00, *b* = 68.23, *c* = 105.80 Å and *a* = 63.06, *b* = 71.80, *c* = 100.23 Å, respectively. The crystals diffracted to resolutions of 2.22 Å (apo) and 2.05 Å (nucleotide bound) using synchrotron radiation and contained one molecule per asymmetric unit. The structures have been determined using the molecular-replacement method and refinement is being undertaken.

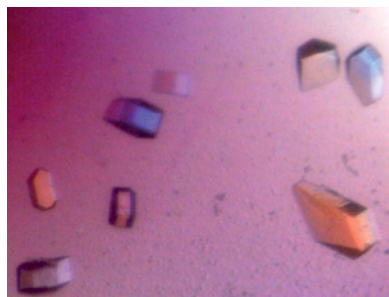
### 1. Introduction

Inositol polyphosphates (IPs) are a new class of signalling molecules that regulate multiple cellular activities, including chromatin remodelling and transcription, mRNA export, telomere-length regulation, RNA editing, exocytosis and translation (Michell, 2008). The levels of these compounds are regulated by different inositol phosphate kinases which phosphorylate the different positions of the inositol ring (Irvine & Schell, 2001).

One of the most prominent and well studied inositol polyphosphates is inositol 1,2,3,4,5,6-hexakisphosphate or phytic acid (IP<sub>6</sub>). Many functions have been suggested for IP<sub>6</sub>, including mRNA export and processing (Michell, 2008). IP<sub>6</sub> acts as a cofactor for different RNA-editing enzymes (Macbeth *et al.*, 2005) and is also required for the production of inositol pyrophosphates (Saiardi *et al.*, 2000), which are involved in apoptosis processes (Nagata *et al.*, 2005). In addition, phytate (the salt of phytic acid) is the major phosphate store in plant seeds; it can lead to malnutrition owing to the chelating power of IP<sub>6</sub> and to environmental phosphorus pollution (Raboy, 2003).

The enzyme in charge of IP<sub>6</sub> synthesis is inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IP<sub>5</sub> 2-K), which employs ATP to phosphorylate the 2-OH axial position of inositol 1,3,4,5,6-pentakisphosphate (Ives *et al.*, 2000). This enzyme is of key importance in mammals; its deletion causes early lethality in mice embryos (Verbsky *et al.*, 2005). The important role of this protein prompted us to crystallize and solve the first structure of an IP<sub>5</sub> 2-K, that from *Arabidopsis thaliana*, in complexes with both substrates and products (Baños-Sanz *et al.*, 2010; González *et al.*, 2010). The protein crystallized in space group *P*<sub>2</sub><sub>1</sub><sub>2</sub><sub>1</sub><sub>2</sub> with two molecules (unit-cell parameters *a* = 58.12, *b* = 113.59, *c* = 139.72 Å; PDB entries 2xal, 2xam, 2xan, 2xar and 2xao) or one molecule (unit-cell parameters *a* = 62.11, *b* = 65.66, *c* = 121.79 Å; PDB entry 4aqq) in the asymmetric unit.

Despite our efforts, crystals of the apo form (without inositol) were not obtained. Study of the apo form is crucial to provide a complete picture of the enzyme and the conformational changes that occur



upon binding of substrates and catalysis. An in-depth analysis of our previous structures suggested that several elements such as Trp129 could be involved in these changes owing to their role in maintaining a closed protein conformation and covering the active site. Therefore, we produced the W129A mutant of IP<sub>5</sub> 2-K by site-directed mutagenesis. Here, we present the expression, purification, crystallization and preliminary X-ray analysis of this mutated IP<sub>5</sub> 2-K. The new crystalline form obtained will be useful in shedding light onto the conformational changes that are undergone by this enzyme.

## 2. Experimental procedures

### 2.1. Protein expression and purification

The W129A mutant was produced by the site-directed mutagenesis technique using 5'-CGTCCGCTAGCGCGTGTTAATGC-3' and 3'-GCATTAACACGCGCTAGCGGACG-5' as oligonucleotides and our previous wild-type construction as the template (Baños-Sanz *et al.*, 2010; González *et al.*, 2010) cloned in pKLSL vector (Angulo *et al.*, 2011). The incorporation of the mutation was assessed by DNA sequencing. Transformation was performed in *Escherichia coli* Rosetta (DE3) pLysS strain; the culture was grown in 2TY medium and incubated with shaking until an OD<sub>600</sub> of 1.0 was reached. Expression was induced with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 16 h at 289 K. The cell pellet was harvested at 3400g and 277 K and stored at 193 K until use.

Pellets were resuspended in buffer A (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM DTT) plus a Complete EDTA-free protease-inhibitor cocktail tablet (Roche) and disrupted with a French press. To eliminate cellular debris, the lysate was clarified at 34 500g for 45 min and filtered with 0.45 μm syringe filters. The filtrated lysate was applied onto a Sepharose CL-6B column equilibrated in buffer A and the protein was eluted in buffer A plus 200 mM lactose. Protein-containing fractions were pooled and diluted threefold with 20 mM Tris-HCl pH 8.0 to reduce the salt concentration, loaded onto a heparin column, washed with buffer B (20 mM Tris-HCl pH 8, 50 mM NaCl, 2 mM DTT) and eluted with a gradient from buffer B to buffer C (20 mM Tris-HCl pH 8, 1000 mM NaCl, 2 mM DTT). The eluted fused protein was cleaved using TEV protease (protease:protein mass ratio 1:80), shaking the sample gently at 277 K overnight. IP<sub>5</sub> 2-K,

LSL and TEV protease were separated by size-exclusion chromatography using a HiLoad 16/600 Superdex 200 column equilibrated in buffer D (20 mM Tris-HCl pH 8.0, 200 mM NaCl, 2 mM DTT). The sample produced an isolated peak that corresponded to an IP<sub>5</sub> 2-K monomer. Pure protein fractions were pooled and loaded into a small Sepharose CL-6B column equilibrated in buffer D to remove residual fusion protein, if present. The final yield of the purification was 14.3 mg per litre of culture. The purity of the sample was confirmed by SDS-PAGE (Fig. 1). The W129A IP<sub>5</sub> 2-K mutant was fully active (data not shown). The final pure sample was concentrated to 10 mg ml<sup>-1</sup> and stored at 193 K.

### 2.2. Protein crystallization

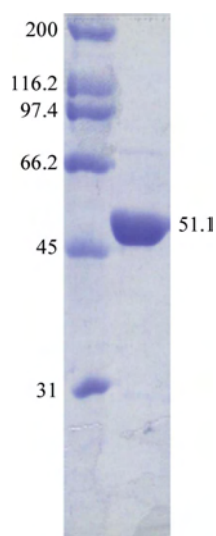
Initial crystallization screening was performed by high-throughput techniques using a NanoDrop robot (Innovadyne Technologies Inc.) in 96-well plates (Innovaplate SD-2 microplates, Innovadyne Technologies Inc.). We tested six commercial screens (Index, Crystal Screen, Crystal Screen 2 and SaltRX from Hampton Research, JB Screen Classic 1–4 from Jena and PACT and JCSG-plus from Qiagen). The sitting-drop vapour-diffusion method was used, mixing 250 nl protein sample with 250 nl precipitant solution. Drops were equilibrated against a reservoir containing 60 μl screening solution. After two weeks, just one condition gave an initial hit: 0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.5, 30% (w/v) PEG 4000 at 293 K (Fig. 2a). This condition was first refined by pH and PEG 4000 grid screening using the sitting-drop vapour-diffusion technique in 24-well Cryschem plates (Hampton Research). The drops (1 μl:1 μl) were equilibrated against a reservoir containing 0.5 ml crystallization condition. We obtained tiny blade-shaped crystals that were not suitable for diffraction studies in conditions consisting of 0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8.0, 32–34% (w/v) PEG 4000 (Fig. 2b). To improve the quality of the crystals, we performed streak-seeding from a seed stock generated from previously obtained crystals. The best crystals, which were suitable for X-ray diffraction, grew in 5 d at 293 K using 0.2 M LiSO<sub>4</sub>, 0.1 M Tris-HCl pH 8, 29% (w/v) PEG 4000 (Fig. 2c).

To obtain nucleotide-bound crystals, we added 2 mM adenosine 5'-(β,γ-imido)triphosphate (AMPPNP; Sigma), a nonhydrolysable analogue of ATP, to the sample. We followed the same strategy for crystallization as was used for the unbound protein. Crystals were obtained using 0.5 M ammonium sulfate, 0.1 M sodium citrate tribasic dihydrate pH 5.6, 1.0 M LiSO<sub>4</sub>. The optimization process was similar as before and a seeding procedure again yielded the best crystals, which were obtained using 0.5 M ammonium sulfate, 0.1 M sodium citrate tribasic dihydrate pH 5.6, 0.8 M LiSO<sub>4</sub>.

### 2.3. Data collection and analysis

The W129A mutant crystals were cryoprotected by gradually increasing the PEG content from 29 to 35% before flash-cooling to 100 K in liquid nitrogen. A complete data set was collected from a W129A IP<sub>5</sub> 2-K crystal at 100 K using an ADSC Q315r X-ray detector on the ID14-4 beamline at the European Synchrotron Radiation Facility (ESRF; Grenoble, France; Fig. 3). AMPPNP-bound crystals were cryoprotected using the crystallization solution plus 20% glycerol and flash-cooled to 100 K. Data for this sample were collected at 100 K using a MAR Mosaic 225 detector on the ID23-2 beamline at the ESRF.

Both data sets were processed using *iMOSFLM* (Battye *et al.*, 2011) and scaled with *SCALA* (Evans, 2006) from the *CCP4* suite (Winn *et al.*, 2011). Data-collection and processing statistics are shown in Table 1.



**Figure 1**  
SDS-PAGE gel of purified W129A IP<sub>5</sub> 2-K. Molecular weights are indicated in kDa.

### 3. Results and discussion

In order to capture a different conformation of IP<sub>5</sub> 2-K by protein crystallography, we produced an IP<sub>5</sub> 2-K point mutant (W129A) by site-directed mutagenesis. The W129A mutant expressed very well, yielding a similar amount of sample as the wild-type protein (Baños-Sanz *et al.*, 2010). However, W129A samples seemed to be more stable than the wild type, with less tendency towards aggregation upon concentration. Wild-type IP<sub>5</sub> 2-K did not crystallize in the absence of inositol, but the introduction of the W129A mutation into the protein allowed us to obtain crystals of unbound and AMPPNP-bound forms.

The diffraction data indicated that the W129A mutant crystals belonged to space group  $P2_12_12$ , with unit-cell parameters  $a = 66.00$ ,  $b = 68.23$ ,  $c = 105.80$  Å. The space group was the same for the AMPPNP-bound W129A mutant crystal, with unit-cell parameters  $a = 63.06$ ,  $b = 71.80$ ,  $c = 100.23$  Å. The calculated values of the Matthews coefficient,  $2.34$  Å<sup>3</sup> Da<sup>-1</sup> for W129A mutant crystals and  $2.24$  Å<sup>3</sup> Da<sup>-1</sup> for AMPPNP-bound W129A mutant crystals, indicated the presence of one protein molecule per asymmetric unit, with a solvent content of 47.4 and 45.0%, respectively. The structure of the W129A mutant was solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007) with our previous structure of IP<sub>5</sub> 2-K bound to IP<sub>5</sub> and AMPPNP (PDB entry 2xan; González *et al.*, 2010) as a model.

**Table 1**

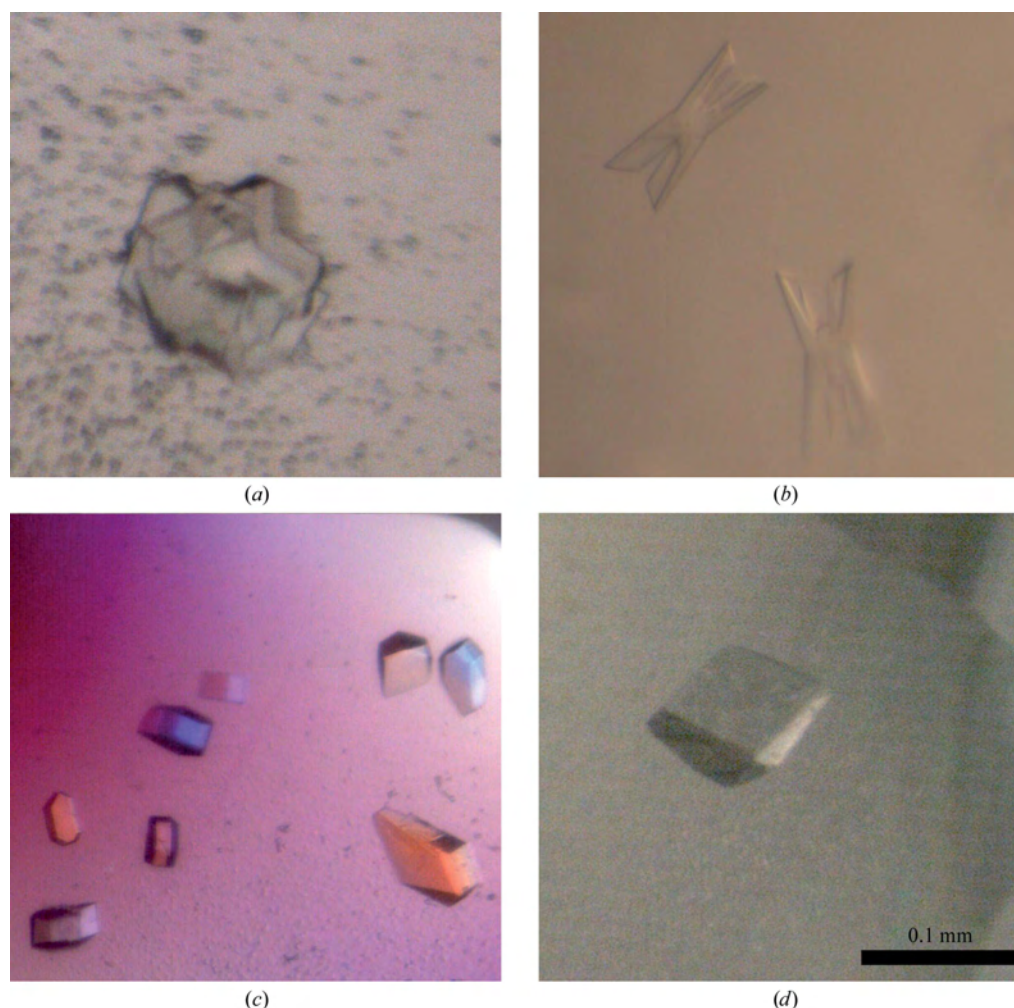
Data-collection and analysis statistics.

Values in parentheses are for the highest resolution shell.

	W129A IP <sub>5</sub> 2-K	W129A IP <sub>5</sub> 2-K-AMPPNP
Temperature (K)	100	100
Wavelength (Å)	0.9794	1.00
Source	ESRF	ESRF
Beamline	ID14-4	ID23-2
Space group	$P2_12_12$	$P2_12_12$
Unit-cell parameters (Å)	$a = 66.00$ , $b = 68.23$ , $c = 105.80$	$a = 63.06$ , $b = 71.80$ , $c = 100.23$
Resolution range (Å)	47.44–2.25 (2.37–2.25)	53.37–2.05 (2.16–2.05)
Unique reflections	169757 (24883)	207910 (30454)
Completeness (%)	99.6 (100)	100 (100)
Multiplicity	7.3 (7.4)	7.1 (7.3)
$R_{\text{merge}}^{\dagger}$ (%)	7.1 (45.9)	9.8 (39.8)
Mean $I/\sigma(I)$	15.9 (4.2)	15.0 (4.9)
Matthews coefficient (Å <sup>3</sup> Da <sup>-1</sup> )	2.34	2.24
Solvent content (%)	47.41	45.03
No. of molecules in asymmetric unit	1	1
Wilson $B$ factor (Å <sup>2</sup> )	47.40	22.00

$\dagger R_{\text{merge}}(I) = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the  $i$ th observed amplitude of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the mean amplitude for measurements of reflection  $hkl$ .

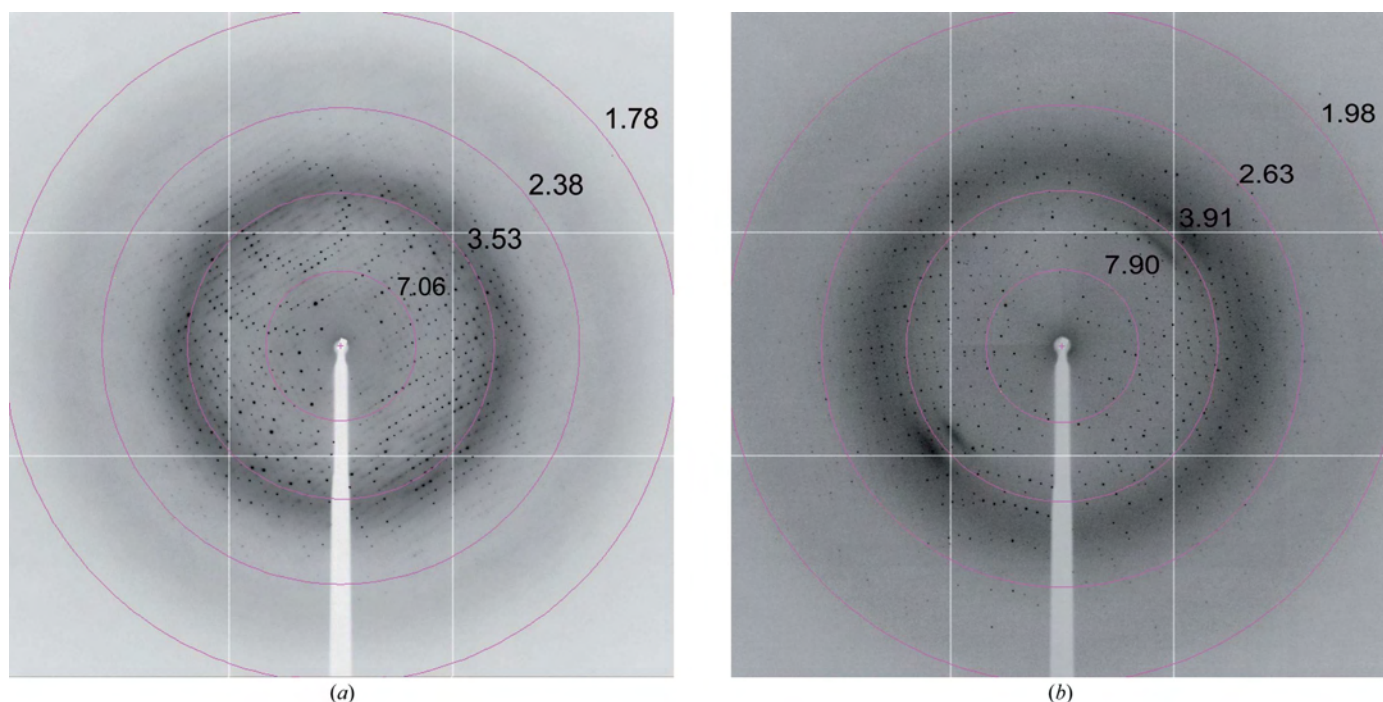
Our initial attempts to solve the structure using this model were unsuccessful; therefore, we defined two ensembles for the search



**Figure 2**

(a) Initial crystals of W129A IP<sub>5</sub> 2-K. (b) W129A IP<sub>5</sub> 2-K crystals obtained in refined conditions without seeding [0.2 M LiSO<sub>4</sub>, 0.1 M Tris–HCl pH 8.0, 33% (w/v) PEG 4000]. (c) W129A IP<sub>5</sub> 2-K crystals obtained after seeding [0.2 M LiSO<sub>4</sub>, 0.1 M Tris–HCl pH 8.0, 29% (w/v) PEG 4000]. (d) AMPPNP-bound W129A IP<sub>5</sub> 2-K crystals (0.5 M ammonium sulfate, 0.1 M sodium citrate tribasic dihydrate pH 5.6, 0.8 M LiSO<sub>4</sub>).





**Figure 3**

X-ray diffraction patterns for (a) apo-form W129A IP<sub>5</sub> 2-K crystals and (b) AMPPNP-bound W129A IP<sub>5</sub> 2-K crystals obtained using synchrotron radiation. The resolution of the rings (in Å) is labelled.

consisting of the N-terminal domain (residues 3–150) and the C-terminal domain (residues 160–437). Using this approach, we found a unique solution with a log-likelihood gain (LLG) of 1456.47 and a translation-function Z score (TFZ) of 23.4. The same approach was carried out to solve the AMPPNP-bound structure. The necessity of separating the protein into two domains to obtain a good solution probably indicates the existence of a conformational change between the inositide-bound form and the unbound form, as expected.

This is a good example of how a simple mutation can alter the crystallization behaviour of a protein. In this case, Trp129 is involved in both crystal packing and active-site closure and its mutation has probably altered both features. Structural refinement is in progress and will allow this hypothesis to be checked and the conformational changes of IP<sub>5</sub> 2-K upon substrate binding to be characterized.

We thank María E. Álvarez-Cao for technical assistance. We also thank the ESRF for providing beamtime and the ID14-4 and ID23-2 staff for providing assistance with data collection. JIB-S is supported by an FPU fellowship (AP2008-00916) from the Ministerio de Educación. This work was supported by grants from Ministerio de Ciencia e Innovación (BFU2008-02897/BMC and BFU2011-24982).

## References

- Angulo, I., Acebrón, I., de las Rivas, B., Muñoz, R., Rodríguez-Crespo, I., Menéndez, M., García, P., Tateno, H., Goldstein, I. J., Pérez-Agote, B. & Mancheño, J. M. (2011). *Glycobiology*, **21**, 1349–1361.
- Baños-Sanz, J. I., Villate, M., Sanz-Aparicio, J., Brearley, C. A. & González, B. (2010). *Acta Cryst. F* **66**, 102–106.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst. D* **67**, 271–281.
- Evans, P. (2006). *Acta Cryst. D* **62**, 72–82.
- González, B., Baños-Sanz, J. I., Villate, M., Brearley, C. A. & Sanz-Aparicio, J. (2010). *Proc. Natl Acad. Sci. USA*, **107**, 9608–9613.
- Irvine, R. F. & Schell, M. J. (2001). *Nature Rev. Mol. Cell Biol.* **2**, 327–338.
- Ives, E. B., Nichols, J., Wenthe, S. R. & York, J. D. (2000). *J. Biol. Chem.* **275**, 36575–36583.
- Macbeth, M. R., Schubert, H. L., Vandemark, A. P., Lingam, A. T., Hill, C. P. & Bass, B. L. (2005). *Science*, **309**, 1534–1539.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Michell, R. H. (2008). *Nature Rev. Mol. Cell Biol.* **9**, 151–161.
- Nagata, E., Luo, H. R., Saiardi, A., Bae, B. I., Suzuki, N. & Snyder, S. H. (2005). *J. Biol. Chem.* **280**, 1634–1640.
- Raboy, V. (2003). *Phytochemistry*, **64**, 1033–1043.
- Saiardi, A., Caffrey, J. J., Snyder, S. H. & Shears, S. B. (2000). *J. Biol. Chem.* **275**, 24686–24692.
- Winn, M. D. *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Verbsky, J., Lavine, K. & Majerus, P. W. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 8448–8453.

# Conformational Changes in Inositol 1,3,4,5,6-Pentakisphosphate 2-Kinase upon Substrate Binding

## ROLE OF N-TERMINAL LOBE AND ENANTIOMERIC SUBSTRATE PREFERENCE\*

Received for publication, March 22, 2012, and in revised form, June 25, 2012 Published, JBC Papers in Press, June 28, 2012, DOI 10.1074/jbc.M112.363671

José Ignacio Baños-Sanz<sup>‡1</sup>, Julia Sanz-Aparicio<sup>‡</sup>, Hayley Whitfield<sup>§2</sup>, Chris Hamilton<sup>¶</sup>, Charles A. Brearley<sup>§3</sup>, and Beatriz González<sup>‡4</sup>

From the <sup>‡</sup>Departamento de Cristalografía y Biología Estructural, Instituto de Química-Física "Rocasolano," CSIC, Serrano 119, 28006-Madrid, Spain and the <sup>§</sup>School of Biological Sciences and <sup>¶</sup>School of Pharmacy, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

**Background:** IP<sub>5</sub> 2-K is essential for higher inositide metabolism and signaling.

**Results:** Crystal structures and biochemical data reveal different IP<sub>5</sub> 2-K conformational states.

**Conclusion:** IP<sub>5</sub> 2-K undergoes conformational changes upon nucleotide and inositide binding, with the N-lobe being essential for substrate recognition and IP<sub>4</sub> enantiomer selection.

**Significance:** Understanding the determinants of enzyme function and substrate specificity will enable rational design of inhibitors.

Inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IP<sub>5</sub> 2-K) catalyzes the synthesis of inositol 1,2,3,4,5,6-hexakisphosphate from ATP and IP<sub>5</sub>. Inositol 1,2,3,4,5,6-hexakisphosphate is implicated in crucial processes such as mRNA export, DNA editing, and phosphorus storage in plants. We previously solved the first structure of an IP<sub>5</sub> 2-K, which shed light on aspects of substrate recognition. However, failure of IP<sub>5</sub> 2-K to crystallize in the absence of inositide prompted us to study putative conformational changes upon substrate binding. We have made mutations to residues on a region of the protein that produces a clasp over the active site. A W129A mutant allowed us to capture IP<sub>5</sub> 2-K in its different conformations by crystallography. Thus, the IP<sub>5</sub> 2-K apo-form structure displays an open conformation, whereas the nucleotide-bound form shows a half-closed conformation, in contrast to the inositide-bound form obtained previously in a closed conformation. Both nucleotide and inositide binding produce large conformational changes that can be understood as two rigid domain movements, although local changes were also observed. Changes in intrinsic fluorescence upon nucleotide and inositide binding are in agreement with the crystallographic findings. Our work suggests that the clasp might be involved in enzyme kinetics, with the N-terminal lobe being essential for inositide binding and subsequent conformational changes. We also show how IP<sub>5</sub> 2-K discriminates between inositol 1,3,4,5-tetrakisphosphate and 3,4,5,6-tetrakis-

phosphate enantiomers and that substrate preference can be manipulated by Arg<sup>130</sup> mutation. Altogether, these results provide a framework for rational design of specific inhibitors with potential applications as biological tools for *in vivo* studies, which could assist in the identification of novel roles for IP<sub>5</sub> 2-K in mammals.

Inositol phosphates (IPs,<sup>5</sup> inositides) are a group of molecules with diverse roles in cell signaling (1). They present crucial roles in DNA reparation and editing, endocytosis, vesicle trafficking, ionic channel regulation, and control of telomere length (2). There are at least 30 different IPs in mammal cells, and their metabolism is dominated by receptor-coupled activation of phospholipase C, which cleaves phosphatidylinositol 4,5-bisphosphate to produce diacylglycerol and inositol 1,4,5-trisphosphate (IP<sub>3</sub>), a starting point for the synthesis of many IPs (2). The levels of these molecules are regulated by families of inositol phosphate kinases or phosphatases that phosphorylate/dephosphorylate the different positions of the inositol ring.

This work is focused on the study of one of those kinases, the inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IP<sub>5</sub> 2-K, IPK1), an enzyme that catalyzes the synthesis of a crucial compound, inositol 1,2,3,4,5,6-hexakisphosphate (IP<sub>6</sub>), by phosphorylation of the axial 2-OH of IP<sub>5</sub>. IP<sub>5</sub> 2-K has been cloned and characterized from many organisms, from yeast to humans (3–8). *Ipk1* gene disruption in mice results in early lethality of the embryos (9). IP<sub>6</sub> participates in mRNA export (10), regulation of chromatin state (11), developmental processes (12), and apo-

\* This work was supported in part by Grants BFU200-02897/BMC and BFU2011-24982 from the Ministerio de Ciencia e Innovación.

The atomic coordinates and structure factors (codes 4axc, 4axd, 4axe, and 4axf) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

<sup>1</sup> Supported by Formación de Profesorado Universitario Fellowship AP2008-00916 from the Ministerio de Educación.

<sup>2</sup> Supported by a Dean's studentship from the Faculty of Science, University of East Anglia.

<sup>3</sup> To whom correspondence may be addressed. E-mail: c.brearley@uea.ac.uk.

<sup>4</sup> To whom correspondence may be addressed. Tel.: 34-91-5619400; Fax: 34-91-564-24-31; E-mail: xbeatriz@iqfr.csic.es.

<sup>5</sup> The abbreviations used are: IP, inositol phosphate; *At*, *A. thaliana*; IP<sub>5</sub> 2-K, inositol 1,3,4,5,6-pentakisphosphate 2-kinase; IP<sub>6</sub>, inositol 1,2,3,4,5,6-hexakisphosphate; IPK, inositol polyphosphate kinase; IP<sub>3</sub>, inositol 1,4,5-trisphosphate; IPmK, inositol polyphosphate multikinase; IP<sub>4</sub> or Ins(1,4,5,6)-P<sub>4</sub>, inositol 1,3,4,5-tetrakisphosphate; Ins(3,4,5,6)P<sub>4</sub>, inositol 3,4,5,6-tetrakisphosphate; Ins(1,3,4,5,6)P<sub>5</sub>, inositol 1,3,4,5,6-pentakisphosphate; AMPNP, adenosine 5'-( $\beta$ , $\gamma$ -imido)triphosphate; r.m.s.d., root mean square deviation.



## IP<sub>5</sub> 2-K Conformational Changes upon Substrate Binding

ptosis (13). In addition, IP<sub>6</sub> serves as substrate for the synthesis of diphosphoinositol polyphosphates (inositol pyrophosphates), emergent molecules with multiple functions (14). In plants, IP<sub>6</sub> has been shown to be involved in the maintenance of basal resistance to plant pathogens (15) and represents a reserve of phosphorus in storage tissues. The high phosphorus content of seeds is a cause of major problems for both human health and the environment (16). In developing countries, grain-based diets provide an excess of IP<sub>6</sub>, which exacerbate iron and zinc malnutrition due to the potent metal-chelating properties of IP<sub>6</sub>. In addition, monogastric animals are unable to digest IP<sub>6</sub>, thus excreting it which leads to water eutrophication (17). Consequently, there is a high demand for the development of low phytate seeds (18–20).

IP<sub>5</sub> 2-K belongs to the inositol polyphosphate kinases (IPKs) family, which comprises three other subfamilies as follows: IP<sub>3</sub> 3-kinases, IP multikinases (IPmKs), and IP<sub>6</sub> kinases (2). The first structure of an IPK, the IP<sub>3</sub> 3-K isoform A (21, 22), became available in 2004 and was followed by the structure of yeast IPmK (23). More recently, we reported the first structure of an IP<sub>5</sub> 2-kinase (24) from *Arabidopsis thaliana*, which showed, unexpectedly, that this enzyme belongs to the IPK family, albeit the most distant member. IP<sub>5</sub> 2-Ks, as well as other IPKs, conserve key features with protein kinases (PKs); they fold in two lobes (N- and C-lobes), connected by a hinge, and the N-terminal lobe core conserves an  $\alpha + \beta$  fold. In addition, the essential features of ATP recognition between both lobes are conserved. However, the C-lobe of IPKs is not conserved with PKs, and the inositide-binding site, located in a separated portion of this C-lobe, is characteristic for each enzyme within the IPK family. In particular, IP<sub>5</sub> 2-K shows the most divergent region for inositide binding (referred to as C<sub>IP</sub>-lobe) that is formed by multiple  $\alpha$ -helices folded into a unique large domain that encompasses almost half of the protein (24). There is a different family of inositol phosphate kinases, exemplified by the *Entamoeba histolytica* enzyme (25), that show the ATP-grasp fold. The most recent addition to this family is the diphosphoinositol pentakisphosphate kinase subfamily, the structure of which was recently solved (26). Although these enzymes display a different fold from IPKs, both share some essential features of catalysis (26). Representatives of all the inositol phosphate kinases mentioned, except IP<sub>6</sub> Ks, have been crystallized and many of them in presence of substrates or products, yielding complete or partial clues about their catalytic mechanism.

The structural analysis of IP<sub>5</sub> 2-K offers a better understanding of its function at the molecular level. The reported complexes of IP<sub>5</sub> 2-K with inositide, both substrate (IP<sub>5</sub>) and product (IP<sub>6</sub>), and the ternary complexes with substrates (IP<sub>5</sub>/AMPPNP) and products (IP<sub>6</sub>/ADP) allowed us to identify key residues defining inositide- and nucleotide-binding sites, as well as crucial elements in the enzymatic mechanism (24). However, there are several aspects that remain to be understood. In particular, all the complexes display a common conformation of the active site that does not correlate with different changes observed in the fluorescence spectra upon substrate binding. We attributed these changes to putative conformational changes that could not be fully characterized because the inositide presence was a requirement for the crys-

tallization of *At*IP<sub>5</sub> 2-K samples (24, 27); consequently, its apo-form remained elusive. In this work, we have followed a strategy for growing crystals from IP<sub>5</sub> 2-K in its free state that was also applied to obtain complexes with nucleotide, always in absence of inositide. The results provide a complete picture of the conformational changes produced upon inositide and/or nucleotide binding that allowed full identification of the molecular landmarks responsible for IP<sub>5</sub> 2-K function. In agreement with the structural studies, fluorescence analysis shows significant IP<sub>5</sub> 2-K changes upon substrate binding. In addition, the exact role of the IP<sub>5</sub> 2-K N-lobe in protein conformational changes and substrate specificity has been also investigated. For this purpose, we have explored discrimination between enantiomers of two IP<sub>4</sub>s, Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub>,<sup>6</sup> that have been reported to be substrates of plant IP<sub>5</sub> 2-K *in vitro*, and we solved the structure of IP<sub>5</sub> 2-K in ternary complex with ATP analog and Ins(3,4,5,6)P<sub>4</sub>. Intriguingly, despite the wealth of literature on Ins(1,4,5,6)P<sub>4</sub> metabolism in yeasts and Ins(1,4,5,6)P<sub>4</sub>/Ins(3,4,5,6)P<sub>4</sub> metabolism in animals, neither of these two isomers are considered physiological substrates of IP<sub>5</sub> 2-K enzymes in these kingdoms. Our experiments provide a structural basis for such consideration.

The results presented here, together with our previous work, represent the first structural description of a complete catalytic cycle of IP<sub>5</sub> 2-K enzymes and gives insights into the discrimination of enantiomeric substrates in IP<sub>5</sub> 2-K.

## EXPERIMENTAL PROCEDURES

**Purification, Crystallization, and Data Collection**—Wild type and W129A IP<sub>5</sub> 2-K mutant were produced in *Escherichia coli* Rosetta (DE3) pLysS strain and purified as reported previously (28). The crystallization of W129A-IP<sub>5</sub> 2-K mutant, in its apo- and AMPPNP-bound forms is detailed in Baños-Sanz *et al.* (28). Briefly, to obtain the W129A IP<sub>5</sub>-2K crystals, equal amounts of protein solution (10 mg/ml) and precipitant (29% w/v PEG4000, 0.2 M LiSO<sub>4</sub>, and 0.1 M Tris/HCl, pH 8.5) were mixed and equilibrated against a reservoir containing 0.5 ml of precipitant using the sitting drop technique. Seeding was used to obtain high quality crystals. Crystals appeared in 5 days at 18 °C. For data collection, crystals were cryoprotected using a gradual increase in the PEG content from 29 to 35% of the crystal mother liquor. For AMPPNP-bound IP<sub>5</sub> 2-K, the mutated protein was mixed with 2 mM AMPPNP, and crystals were obtained in 0.5 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 M sodium citrate tribasic, pH 5.6, and 0.8 M Li<sub>2</sub>SO<sub>4</sub> and cryo-protected with this condition plus 20% glycerol. Seeding was again crucial to facilitate crystal growth. The ADP-bound IP<sub>5</sub> 2-K crystals were obtained in the same condition, except the protein was incubated previous to crystallization trials with 2 mM ADP. To obtain wild type crystals in complex with Ins(3,4,5,6)P<sub>4</sub> and AMPPNP, we have followed our protocol previously reported (27). In this case, protein was incubated with 1 mM IP<sub>4</sub> plus 2 mM AMPPNP, crystallized, and cryoprotected under the same

<sup>6</sup> The following nomenclature is used: stereoisomers and enantiomers of inositol phosphates are numbered according to the D-assignment, thus D- and L-Ins(1,4,5,6)P<sub>4</sub> are identified as Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub>, respectively.

TABLE 1

## Crystal data collection, structure determination, and refinement

Data for the outermost shell are given in parentheses. a.u. means asymmetric units.

Crystal	W129A	W129A/AMPPNP	W129A/ADP	IP <sub>5</sub> 2-K/AMPPNP/IP <sub>4</sub>
<b>Data collection</b>				
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell	66.00 68.23 105.80	63.06 71.80 100.23	63.72 71.92 100.62	61.62 65.08 122.49
Wavelength	0.9794 Å	1.0053 Å	0.9334 Å	0.9334 Å
Resolution range	47.44 to 2.25 Å	53.37 to 2.05 Å	39.49 to 2.50 Å	57.47 to 2.93 Å
R <sub>p</sub> <sup>a</sup>	0.028 (0.180)	0.04 (0.158)	0.054 (0.204)	0.059 (0.161)
R <sub>merge</sub> <sup>b</sup>	0.072 (0.517)	0.098 (0.398)	0.105 (0.407)	0.146 (0.394)
Unique reflections	24,181 (3488)	29,280 (4191)	16,575 (2370)	77,497 (10881)
Redundancy	7.3 (7.4)	7.1 (7.3)	4.6 (4.8)	7.0 (6.9)
Completeness	99.6% (100%)	100% (100%)	99.8% (100.0%)	100.0% (100.0%)
Mean $I/\sigma(I)$	15.4 (3.8)	15.0 (4.9)	10.1 (3.6)	13.4 (4.9)
$I/\sigma(I)$	6.1 (1.5)	5.5 (1.9)	5.4 (1.9)	4.9 (1.9)
Wilson $B$ value	47.4 Å <sup>2</sup>	22.0 Å <sup>2</sup>	39.7 Å <sup>2</sup>	39.2 Å <sup>2</sup>
<b>Refinement statistics</b>				
Molecules per a.u.	1	1	1	1
Resolution range	47.44 to 2.25 Å	53.37 to 2.05 Å	39.49 to 2.50 Å	57.47 to 2.93 Å
R <sub>factor</sub> <sup>c</sup>	0.24%	0.20%	0.22%	0.22%
R <sub>free</sub>	0.28%	0.25%	0.28%	0.28%
Protein atoms (non-H)	3151	3458	3391	3260
Ligand atoms (non-H) IP <sub>4</sub> , IP <sub>6</sub> , AMPPNP, ADP	0	31	27	59
Ions SO <sub>4</sub> <sup>2-</sup> , Zn <sup>2+</sup>	21	26	31	1
Waters	74	324	66	0
r.m.s.d.				
Bonds	0.009 Å	0.008 Å	0.005 Å	0.005 Å
Angles	1.69°	1.69°	1.04°	1.06°
Mean $B$ value				
Protein	59.27 Å <sup>2</sup>	24.15 Å <sup>2</sup>	41.10 Å <sup>2</sup>	29.05 Å <sup>2</sup>
Ligands	NA <sup>d</sup>	19.10	32.15	18.34
Waters	30.79	30.45	32.07	NA
Missing residues	1–2, 47–60, 152–161, 334–341, 382–387, 434–451	386, 434–451	158, 159, 334–336, 386, 434–451	47–58, 155–158, 378–386, 437–451
<b>Ramachandran data</b>				
Most favored	92.5%	91.2%	92.4%	92.3%
Disallowed	0.3%	0.5%	0.3%	0.3%

<sup>a</sup>  $R_{pim} = \sum (1/(N-1))^{1/2} \sum |I(h)i - \langle I(h) \rangle| / \sum \langle I(h) \rangle$ .<sup>b</sup>  $R_{merge} = 100 \sum h \sum i |I(h, i) - \langle I(h) \rangle| / \sum h \sum i I(h, i)$ , where the outer sum ( $h$ ) is over the unique reflections and the inner sum ( $i$ ) is over the set of independent observations of each unique reflection.<sup>c</sup>  $R$ -factor =  $\sum (|F_{obs} - F_{calc}| / \sum |F_{obs}|)$  ( $R_{free}$  is equivalent to  $R$ -factor for a randomly selected 5% subset of reflections not used in structure refinement).<sup>d</sup> NA means not applicable.

conditions as for the other wild type crystal complexes (24). Diffraction data for all the crystals was collected using synchrotron radiation from European Synchrotron Radiation Facility (Grenoble, France). All the data were processed using iMosflm (29) and merged with Scala (30) from the CCP4 package (31). A summary of data collection statistics is shown in Table 1.

**Structure Solution and Refinement**—All the structures were solved by Molecular Replacement with Phaser program (32), using as a search model our previous wild type structure (Protein Data Bank code 2xan) (24) and dividing the model in two ensembles for the W129A mutant structures, as detailed previously (28). Several rounds of refinement with Refmac5 (33) were alternated with model building with the program Coot (34). Final refinement parameters are reported in Table 1. In particular, the electron density maps of the W129A-IP<sub>5</sub> 2-K crystal did not allow building of regions consisting of residues 46–59 in hinge 1 and part of hinge 3, whereas the W129A-IP<sub>5</sub> 2-K-nucleotide complex density maps allowed almost complete building of the model. The average  $B$ -factors of both crystals show significant differences, which is coherent with a high flexibility in the open form of the free enzyme, and the stabilization exerted by AMPPNP binding. The ADP-bound crystals

present higher  $B$ -factors, revealing the importance of Py in the stabilization. Stereochemistry of the model was checked using PROCHECK (35). Protein pictures were performed with PyMOL (36). Protein motion along IP<sub>5</sub> 2-K different states has been analyzed manually and automatically using DymDOM server (37).

**Mutant Production and Purification**—The mutants was performed using the QuikChange protocol, employing the following pairs of oligonucleotides: 5'-GATTCGTATACAGAAGGCTGCGAGGAATGATAAAGCAATC-3' and 5'-GATTGCTTTATCATTCCTCGCAGCCTTCTGTATACGAATC-3' for R45A mutation; 5'-CGTCCGCTAGCGCGTGTAAATGC-3' and 5'-GCATTAACACGCGCTAGCGGACG-3' for W129A mutation; 5'-CGTCCGCTAGTGGCTGTAAATGC-3' and 5'-GCATTAACAGCCACTAGCGGACG-3' for W129V/R130A mutation; 5'-CGTCCGCTAGTGGCTGTAAATGC-3' and 5'-GCAGGCGATACCGCACAATTACG-3' for R130A mutation; and 5'-CCCGGTGCTACCACCCCTAATATGAGAGAACC-3' and 5'-GGGGGTGGTAGCACCAGGAGAACAGCCCC-3' for ΔS253/ΔE255 deletions. Previous construction of wild type fused to lectin from *Letiporus sulphureus* was used as template (24). The incorporation of the mutations was assessed by DNA sequencing. Mutants were grown in *E. coli*

## IP<sub>5</sub> 2-K Conformational Changes upon Substrate Binding

Rosetta (DE3) pLysS strain as wild type protein. Cell pellets were resuspended in buffer A (20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM DTT plus a Complete EDTA free protease inhibitor mixture (Roche Applied Science)) and disrupted with a French press. The filtrated cell lysate was applied onto a Sepharose CL-4B column equilibrated previously with buffer A, and the protein was eluted with buffer A plus 200 mM lactose. Protein-containing fractions were pooled and diluted 3-fold with 20 mM Tris, pH 8, to reduce the salt concentration, loaded onto a heparin column, washed with buffer B (20 mM Tris, pH 8, 50 mM NaCl, 2 mM DTT), and eluted with a gradient between buffer B and C (20 mM Tris, pH 8, 1 M NaCl, 2 mM DTT). The eluted fused protein was cleaved by tobacco etch virus protease (protease/protein mass ratio 1:80), shaking the sample gently at 4 °C overnight. To remove the lectin of *L. sulphureus* tag, we used another Sepharose CL-4B column. Mutants were concentrated between 0.5 and 1 mg/ml and stored at –80 °C until use.

**Enzyme Assays**—Kinase assays were performed in 0.2-ml volumes on a Molecular Devices SpectraMax M5 Plate reader using an assay similar to that described previously (38, 39). The assay mixture contained 20 mM Hepes, pH 7.5, 1 mM MgCl<sub>2</sub>, 0.1 mM ATP, 2 mM P-enolpyruvate, 0.15 mM NADH and 7.5 units ml<sup>–1</sup> lactate dehydrogenase, 15 units ml<sup>–1</sup> PK, 5 μM inositol phosphate. Reactions were started by addition of the above assay mixture to enzyme and 100 or 500 ng of native protein or mutant as specified in the text. Assays were run for 15 min at 25 °C and initial linear rate measurements (ΔOD min<sup>–1</sup>) were determined using the SoftMax Pro software of the plate reader: initial rates were linear for 3–9 min. Assays were typically performed with 4–5 replicate wells and were repeated no less than three times.

**Kinetic Analysis of Native and W129A Mutant**—Kinase assays were performed with 100 ng of native or W129A protein and 0.3–10 μM Ins(1,3,4,5,6)P<sub>5</sub>. Assays were performed with 4–5 replicate wells and repeated three times.  $K_m$ ,  $v_{app}$  and  $V_{max, app}$  values were determined by nonlinear least squares regression of a plot of  $V_{app}$  versus substrate concentration fitted to the Michaelis-Menten equation in GraFit Version 5 (Erithacus Software Ltd.).

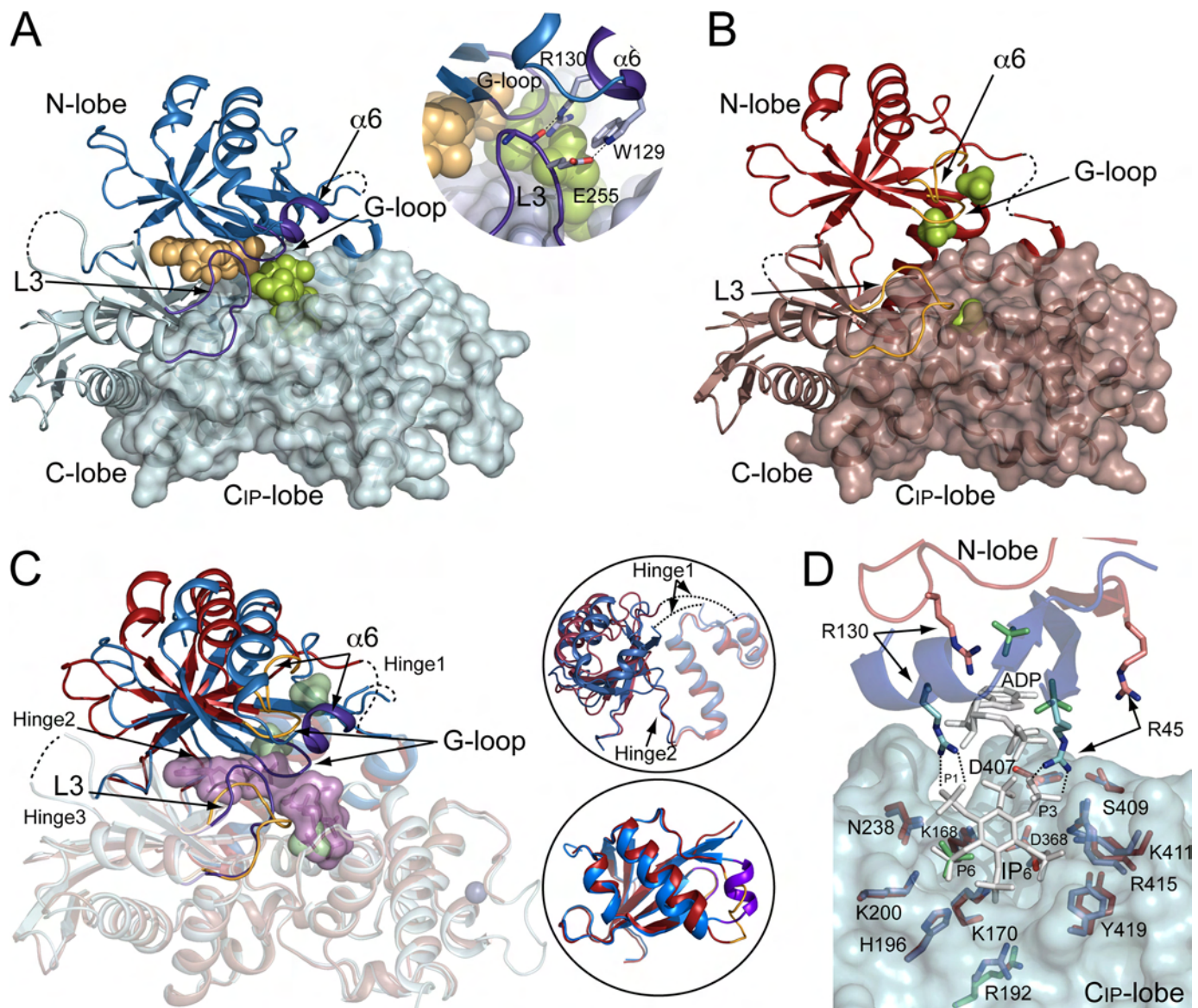
**Fluorescence Assays**—Substrate binding to native and W129A protein was followed at 25 °C by monitoring changes in intrinsic fluorescence recorded with an excitation wavelength of 295 nm, bandwidth 5 nm, and emission from 310 to 380 nm, bandwidth 10 nm, on a Cary Eclipse fluorescence spectrometer. The photomultiplier voltage was set to 800 V. Protein (14.5 μg) was suspended in 1 ml of 20 mM Hepes, pH 7.5, 1 mM MgCl<sub>2</sub> in a 10 × 3-mm cuvette. Inositol 1,3,4,5,6-pentakisphosphate or AMPPNP was added from 0.1 and 5 mM stocks, respectively. The total volume of ligand added was less than 1% of the sample volume. Data were exported in GraFit Version 5 (Erithacus Software Ltd.). Plots of fractional change in fluorescence ( $F_o - F$ )/ $F_o$  versus ligand concentration, determined at 335 nm) were analyzed by nonlinear least squares regression to a single site binding model in GraFit Version 5. Experiments were repeated three times.

## RESULTS

**Structure of IP<sub>5</sub> 2-K Apo-form Represents the Open Conformation**—Previously reported IP<sub>5</sub> 2-K inositide-bound structures, with or without bound nucleotide, show a common conformation that leaves a hollow for ATP binding formed between the N-lobe and C-lobe. The inositide is mainly bound to the C-lobe (at the unique region named C<sub>IP</sub>-lobe), but it also binds to a few N-lobe residues bringing together both lobes (Fig. 1A). We will refer to this state as “closed” or “inositide-bound” form. In this form, the β- and γ- phosphates of the ATP are linked to the G-loop (Glu<sup>18</sup>–Asn<sup>21</sup>), a feature common to PKs, and the active site is enclosed by residues from the structural elements α6 (Leu<sup>128</sup>–Ala<sup>133</sup>) and L3 (Gly<sup>251</sup>–Ser<sup>261</sup>) that interact forming a kind of clasp (Fig. 1A, *zoom*). Because the clasp seems to preclude the nucleotide from entering/leaving the active site, we suspected that the clasp elements might be unconnected in a putative open conformation, and therefore, we expected that L3 would present high mobility. Accordingly, the IP<sub>5</sub> 2-K·IP<sub>5</sub> crystal complex showed some disorder at L3 (Protein Data Bank code 2xao). Moreover, AtIP<sub>5</sub> 2-K exhibits changes in gel filtration elution profile (data not shown) and in tryptophan intrinsic fluorescence (see below) upon substrate binding, indicative of structural changes, which in turn could arise from changes in the environment of Trp<sup>129</sup>. We have therefore undertaken an in-depth study of the clasp role through characterization of several IP<sub>5</sub> 2-K mutants (W129A, W129V/R130A, R130A, and ΔS253/ΔE255) (Table 2). One of these fully active mutants, W129A, allowed us to obtain crystals from the enzyme in its apo-form, *i.e.* in its unbound state. The crystals obtained present a different unit cell, space group, and crystal packing than the inositide-bound structures (Table 1) (28).

The W129A IP<sub>5</sub> 2-K structure shows a similar overall fold to that reported previously for the enzyme, but the N-lobe and C-lobe are farther apart from each other forming a more open cleft (Fig. 1B). We will refer to this form as “open form” or “apo-form.” Superposition of Cas from open and closed (24) conformations gives an r.m.s.d. of 2.50 Å; however, a close inspection reveals that there is a significant motion of two independent portions of the protein (Fig. 1C) as follows: one is formed by most of the N-lobe, which includes residues 6–42 and 103–148 (r.m.s.d. of 1.27 Å superposing 81 residues from each state); and the other is composed of a segment of the N-lobe previously referred to as N-I (24) that includes residues 63–102, and the whole C-lobe (r.m.s.d. of 0.87 Å for 302 residues). This motion upon substrate binding can be described as an 18.4° rotation of one portion of the enzyme, which pivots about three regions of the polypeptidic chain acting as hinges; two of them are the two segments that connect N-I to the remaining part of the N-lobe (hinge 1, 42–63; hinge 2, 102–103) and the third one is the linker between N- and C-lobe (hinge 3: 148–159), a typical hinge common to the family of PKs (Fig. 1C). As expected, these hinges are the most disordered parts in the electron density reflecting its intrinsic flexibility. Apart from this general rigid-body movement, local changes occur particularly at α6 and L3 segments (Fig. 1C, and *bottom*





**FIGURE 1. Structure of IP<sub>5</sub> 2-K open form versus closed form.** A, IP<sub>5</sub> 2-K closed form structure, IP<sub>5</sub> 2-K complex with ADP and IP<sub>6</sub>. The N-lobe and C-lobe are shown as blue and light blue schematics, respectively. The CIP-lobe is shown as surface representation. ADP and IP<sub>6</sub> are shown as orange and green spheres, respectively. Disordered loops are shown as dashed lines. The key elements that close the active site (G-loop,  $\alpha 6$ , and L3) are shown in purple-blue. On the top right, a zoom of the above elements is shown, showing a detail of the interaction between  $\alpha 6$  (N-lobe) and L3 (C-lobe). This interaction forms a clasp, joining both lobes and closing the active site. B, schematic representation of IP<sub>5</sub> 2-K open form as seen in the W129A mutant structure. The N-lobe and C-lobe are represented as dark and light red colors. The CIP-lobe is highlighted with surface representation. Elements involved in covering the active site in the closed form are pointed out and are highlighted in orange color. Three sulfate molecules are represented as green surface. Disordered loops are shown as dashed lines. C, superposition of IP<sub>5</sub> 2-K open (red colors) and closed (blue colors) forms. The conformational changes upon inositol binding are well illustrated when superposing a region containing the C-lobe and segment N-I (shown as transparent schematics). In this superposition, a movement of a large portion of N-lobe is highlighted. Substrates in the closed form are shown as violet transparent surface, and sulfate ions observed in the open form are shown as green surface. The three hinges of the structure are also marked. The top right panel zooms the N-lobes of both forms, where the two regions can be appreciated as follows: a nonrotated (transparent) and a rotated region (solid), both being connected by hinges I and II. The bottom right panel shows a superposition of the rotated regions within the N-lobe (residues 6–42 and 103–148 from N-lobe) showing that additional local changes occur upon inositol binding, mainly at  $\alpha 6$ . D, inositol-binding site differences in open (red) and closed (blue) forms. The superposition made is similar to C. Both N-lobes are shown as schematic representation and the C-lobe as a transparent surface. The inositol ligands are shown as stick representation. No major variation is found in the C-lobe residues involved in inositol binding, and the N-lobe residues are farther apart in the open form, revealing that the inositol site is not fully formed in the open form. ADP and IP<sub>6</sub> in the complex are shown as white sticks, and the three sulfate moieties of open form are shown as green sticks. The G-loop is also involved in inositol binding but is not shown for clarity.

zoom), and their conformation is dependent on the clasp formation.

A detailed comparison of open and closed IP<sub>5</sub> 2-K forms reveals that the CIP-lobe pocket remains unaltered upon inositol binding (Fig. 1D). Therefore, we can assume that the major inositol-binding site is preformed in the apo-form. However,

this site is not fully created in the open form because the N-lobe is too far away to accomplish inositol recognition. This recognition involves interactions with the side chain of residues Arg<sup>45</sup> and Arg<sup>130</sup> and the main chain of the G-loop (Fig. 1D).

A last interesting feature is the presence of three sulfate ions in the crystal at positions similar to those occupied by P<sub>γ</sub> (ATP)

**TABLE 2**

**Activity of wild type and mutant IP<sub>5</sub> 2-K toward Ins(1,3,4,5,6)P<sub>5</sub>**

Assays were performed at 4  $\mu$ M IP<sub>5</sub>, 100  $\mu$ M ATP. Results obtained from initial rate measurements are expressed relative to the activity of wild type enzyme assayed on the same microplate, mean  $\pm$  S.D. of three experiments with 4–5 replicate measurements per experiment.

	Mean	S.D.
Wild-type	100	
R45A	58	10
W129A	121	11
W129V/R130A	21	11
R130A	16	9
$\Delta$ Ser <sup>253</sup> / $\Delta$ Glu <sup>255</sup>	327	36

and by P1 and P6 (inositide), as found in the IP<sub>5</sub> 2-K ternary complex structure (Fig 1, B and D). These sulfate moieties, present in the crystal buffer, mimic the binding positions of the substrate in the open form of IP<sub>5</sub> 2-K.

**IP<sub>5</sub> 2-K Nucleotide-bound Structure Presents a Half-closed Conformation**—In our previous work, we have reported the structure of IP<sub>5</sub> 2-K ternary complexes with either substrates (IP<sub>5</sub> + AMPPNP) or products (IP<sub>6</sub> + ADP) bound at the active site (21, 24). The structure of these ternary (nucleotide + inositide) complexes revealed a common conformation with inositide-IP<sub>5</sub> 2-K binary complexes. However, the permanent presence of the inositide in our crystals precluded the assessment of specific conformational changes produced upon ATP binding. A new full crystal screening allowed us to crystallize W129A IP<sub>5</sub> 2-K with the nucleotide (ATP or ADP) (28). Although these crystals appeared in very different conditions and show different cell parameters with respect to the apo-form, they present the same space group and crystal packing. Superposition of the nucleotide-bound coordinates to those corresponding to the open and closed conformations of IP<sub>5</sub> 2-K shows that the enzyme is in an intermediate state somewhere between the apo-form (r.m.s.d. = 1.47 Å) and the closed form adopted by the ternary complex (r.m.s.d. = 1.71 Å) (Fig. 2A). Therefore, IP<sub>5</sub> 2-K reaches a half-closed conformation upon ATP binding that corresponds to a 10° rotation of the same portion of the N-lobe described above. We refer to this form as the “nucleotide-bound” or “half-closed” form. The ATP-binding site formed is essentially the same as that found in the ternary complexes, with only two significant differences as follows. First, in the absence of inositide there is no evidence of Mg<sup>2+</sup> binding in the electron density map. This could be due to the different lobes of orientation in this form, which conforms a Mg<sup>2+</sup>-binding site slightly distorted as compared with that observed in the closed form. Second, a small tilt in ATP position is observed upon inositide binding with a better fit of the nucleotide into a slightly more closed pocket in the ternary complex (Fig. 2B). Despite this, the nucleotide pocket formed conserves all the essential features of the nucleotide recognition by the protein kinase superfamily, revealing that this conformation is physiologically relevant. In contrast, the C<sub>IP</sub>-lobe remains unaltered upon nucleotide binding, and the N-lobe, although closer, is still too far away to complete the inositide-binding site; consequently, the clasp observed in the closed-form is not formed.

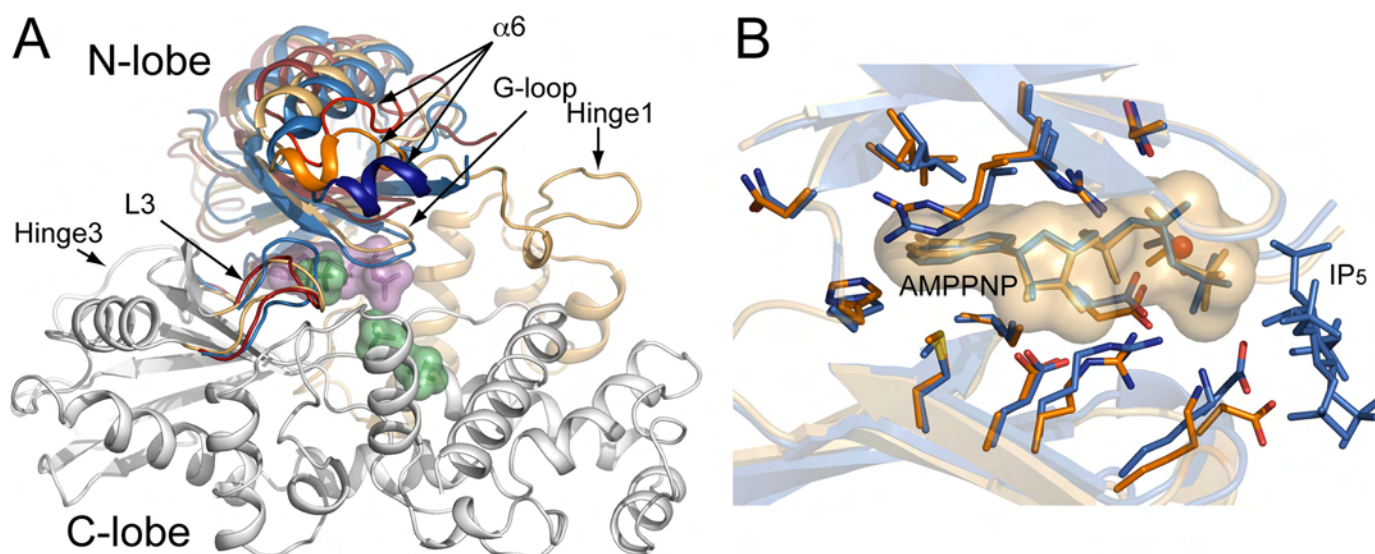
In the half-closed form that we obtained with nucleotide, sulfate ions were found bound in similar positions to that occupied in the closed form by inositide phosphates P1, P6, and P5

(Fig. 2A). These sulfates putatively mimic the binding of inositide in this half-closed form. Finally, unlike what was observed in the open and closed conformations, hinge 1 and hinge 3 are ordered in this form.

**Snapshots of the Conformational Changes**—The crystal structures of the two IP<sub>5</sub> 2-K states here reported, together with our previous work (24), provide more detail of the enzymatic mechanism. Our experimental results reveal essential structural motions associated with nucleotide and inositide binding that not only involve a large movement of two protein portions as rigid bodies but also local changes on the clasp region (Fig. 3). As predicted before, the clasp is fully formed only in the closed state by interaction of L3 (from the C-lobe) and  $\alpha$ 6 (from the N-lobe), through Trp<sup>129</sup>–Glu<sup>255</sup> and Arg<sup>130</sup>–Gly<sup>254</sup> links (Fig. 3A, right). The formation of this clasp requires the rupture of the Glu<sup>255</sup>–Lys<sup>200</sup> ion pair that links L3 to the C<sub>IP</sub> lobe, in both the open and nucleotide-bound conformations (Fig. 3A, left and middle panels). This breakage releases L3 to act as the “lid” that closes the pocket upon inositide binding. The conformation of both  $\alpha$ 6 and L3 changes significantly among the different structural states of IP<sub>5</sub> 2-K (Fig. 3A). In fact, the segment  $\alpha$ 6 is not structured in the open conformation; the helix is fully formed only in the closed state. In addition, L3 is rearranged in the closed state, taking an extended conformation formed by a new pattern of interactions with IP<sub>5</sub> 2-K residues of loop Thr<sup>235</sup>–Asn<sup>239</sup> (L2), where the conserved <sup>237</sup>QNN motif is located (Fig. 3B). Thus, the interaction Asn<sup>238</sup>–Arg<sup>241</sup> observed in the open and half-closed conformation is broken leaving L2 able to interact with L3 through Gln<sup>237</sup>–Thr<sup>257</sup> and Asn<sup>238</sup>–Ser<sup>256</sup> links. Consequently, most of the structural changes observed in IP<sub>5</sub> 2-K are due to inositide binding, but it is worth mentioning that in the nucleotide-bound form the clasp seems partially created through interaction of Arg<sup>130</sup>( $\alpha$ 6) and Gly<sup>252</sup>(L3).

**Clasp Role**—As shown in Table 2, we have measured the activity of several IP<sub>5</sub> 2-K mutants of clasp-contributing residues to assess their contribution to the enzymatic mechanism. Assays were performed by a coupled enzyme assay allowing steady-state measurements. In particular, we have analyzed the effect of single and double substitutions in  $\alpha$ 6 key residues (W129A, R130A, and W129V/R130A) and deletions in L3 ( $\Delta$ Ser<sup>253</sup>/ $\Delta$ Glu<sup>255</sup>). First, the activity of the W129A mutant is uncompromised leading to the conclusion that Trp<sup>129</sup> participation in clasp formation is not critical for enzyme activity. It is worth noting that the bond formed by Trp<sup>129</sup> in the clasp is not the unique determinant for the clasp formation. Second, mutation of Arg<sup>130</sup> results in more than 6-fold decreased activity toward Ins(1,3,4,5,6)P<sub>5</sub> and toward other substrates (see below), and the double mutation W129V/R130A results also in a 5-fold reduced activity, a fact that could be ascribed mainly to the Arg<sup>130</sup> mutation. Arg<sup>130</sup> is involved in both clasp formation and inositide binding through P1 (Figs. 1D and 5); therefore, its contribution may be more critical for enzyme function. Finally, shortening of L3 by two residues ( $\Delta$ Ser<sup>253</sup>/ $\Delta$ Glu<sup>255</sup>) results in an IP<sub>5</sub> 2-K mutant 3-fold more active than the native form. Modeling this shorter loop in the IP<sub>5</sub> 2-K structure suggests that deletion of these two residues precludes clasp formation. We suggest that the increased activity may be attributed to an improved kinetics of the enzyme, perhaps through removal of





**FIGURE 2. Conformational changes upon nucleotide binding.** *A*, comparison of the three IP<sub>5</sub> 2-K obtained conformations (open, half-closed, and closed forms). The superposition has been made on the C-lobes and segment N-I as before. The IP<sub>5</sub> 2-K nucleotide-bound form is shown as *orange* (N-lobe) and *white* (C-lobe) schematics. The nucleotide and the sulfate moieties found in this form are shown as *transparent violet* surface and *green* surface, respectively. As the C-lobes superpose very well in the three cases, only the rotated N-lobe region and L3 have been shown for the open (*red*) and closed (*blue*) forms. This picture highlights the sequential structural changes upon nucleotide and inositol binding, as well as the clasp elements variation. *B*, comparison between the nucleotide-binding site as shown in the half-closed (*orange*) and closed (*blue*) forms. The superposition has been made by matching the ATP molecules. The ATP-binding site is very similar in both structures, although it is more constrained in the closed form. Magnesium found in the closed form is shown as *red sphere*.

the lid, that would decrease the energetic barrier of the process, for instance through a more favorable release of products.

As mentioned above, we have shown that W129A-IP<sub>5</sub> 2-K retains as much activity as the native enzyme (Table 2). We have performed kinetic measurements for native and W129A IP<sub>5</sub> 2-K (Fig. 4 and Table 3). Our results show that, although there are no dramatic changes, the native IP<sub>5</sub> 2-K presents ~3-fold greater affinity for the inositol substrate ( $K_{m,app}$   $3.7 \pm 1.0 \mu\text{M}$ ) than the mutated enzyme ( $13.7 \pm 3.0 \mu\text{M}$ ). However,  $V_{max,app}$  of W129A mutant ( $0.95 \pm 0.16 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ) is 2-fold higher than that of the native enzyme ( $0.54 \pm 0.07 \mu\text{mol min}^{-1} \text{mg}^{-1}$ ), telling us that the mutant is more efficient. The removal of the Trp<sup>129</sup> side chain probably affects the inositol binding pocket formation, as it is next to Arg<sup>130</sup>. In turn, Trp<sup>129</sup> removal likely weakens clasp formation producing a more efficient enzyme, which correlates with our previous suggestion that lid removal could be favoring the product release.

**Role of the N-lobe in Inositol Binding and Specificity**—Inositol 1,3,4,5,6-pentakisphosphate is bound to the N-lobe through phosphates P1 (Arg<sup>130</sup>) and P3 (Arg<sup>45</sup> and G-loop main chain) (Fig. 1D). To assess the importance of P1 and P3 to inositol binding and ensuing conformational change, we made several attempts to crystallize IP<sub>5</sub> 2-K in presence of the enantiomers Ins(3,4,5,6)P<sub>4</sub> and Ins(1,4,5,6)P<sub>4</sub>, which are expected to bind in the same orientation as Ins(1,3,4,5,6)P<sub>5</sub>. The presence of a plane of symmetry along the C2-C5 axis of Ins(1,3,4,5,6)P<sub>5</sub> allows, in comparison with the two IP<sub>4</sub>s, an assessment of the contribution of residues coordinating the enantiotopic 1- and 3-phosphates.

We obtained crystals of IP<sub>5</sub> 2-K in ternary complex with Ins(3,4,5,6)P<sub>4</sub> and AMPPNP. These crystals turned out to be similar to the previously solved ternary complexes and binary complexes of IP<sub>5</sub> 2-K and inositol (r.m.s.d. 0.425 Å) (24) and

confirm that Ins(3,4,5,6)P<sub>4</sub> binds in the same orientation as Ins(1,3,4,5,6)P<sub>5</sub> (Fig. 5). The most interesting feature of the new ternary complex is that Arg<sup>130</sup> is perfectly oriented to potentially bind the absent P1, and it maintains interaction with L3 forming the clasp as in IP<sub>5</sub>-bound structures. This result suggests that Arg<sup>130</sup> participation in P1 binding and clasp formation represents two separate facets of IP<sub>5</sub> 2-K function. As mentioned before, the R130A mutant presents reduced but significant IP<sub>5</sub> 2-K activity, suggesting that although Arg<sup>130</sup> is an important residue, it is not essential for catalysis. In parallel, as we were unsuccessful in attempts to grow crystals in the presence of inositol lacking P3 (Ins(1,4,5,6)P<sub>4</sub>), we have generated an R45A IP<sub>5</sub> 2-K mutant to remove one of P3 binding residues (Figs. 1D and 5). The fact that R45A retains appreciable activity against Ins(1,3,4,5,6)P<sub>5</sub> suggests that Arg<sup>45</sup>, although not essential, makes some contribution to inositol binding. As P3 of both Ins(1,3,4,5,6)P<sub>5</sub> and Ins(3,4,5,6)P<sub>4</sub> are bound to the protein G-loop, it seems likely that the G-loop plays a main role in inositol binding and in subsequent conformational changes.

Because there remains conjecture concerning the relative contribution of different pathways to IP<sub>6</sub> synthesis in plants (19), and because Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub> have been reported to be substrates *in vitro* of plant enzymes (6, 18), we have compared the activity of IP<sub>5</sub> 2-K toward Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub> for native protein and for a range of mutants in N-lobe residues. Our assays were performed at a fixed concentration of inositol phosphate (5  $\mu\text{M}$ ) and ATP (100  $\mu\text{M}$ ) (Table 4). Ins(3,4,5,6)P<sub>4</sub> and Ins(1,4,5,6)P<sub>4</sub> are equally active substrates for IP<sub>5</sub> 2-K. The ratio of activity against Ins(1,4,5,6)P<sub>4</sub>/Ins(3,4,5,6)P<sub>4</sub> is increased significantly by the R45A mutation and reduced slightly by W129A mutation. However, and remarkably, substitution involving Arg<sup>130</sup> (R130A) increase



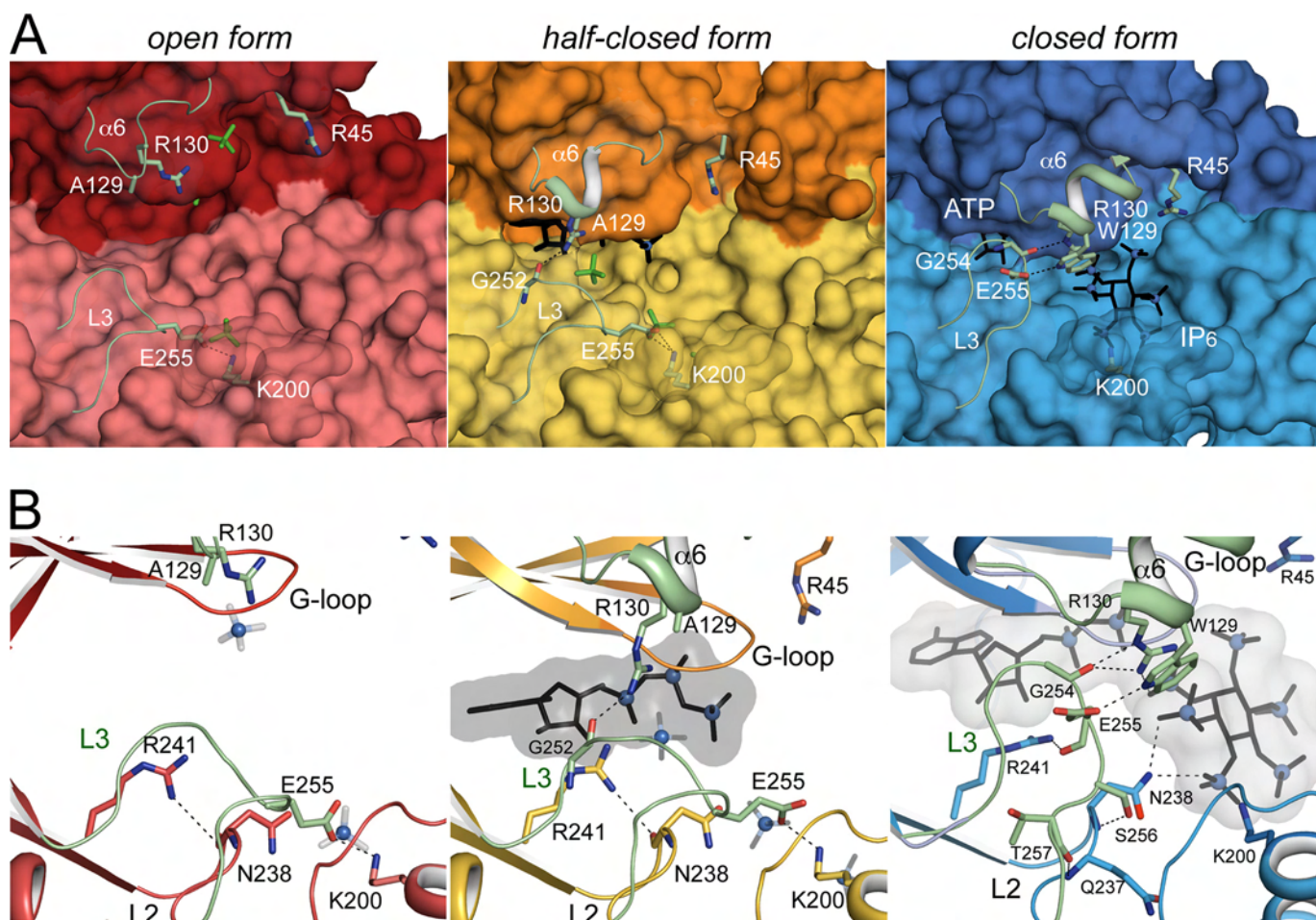


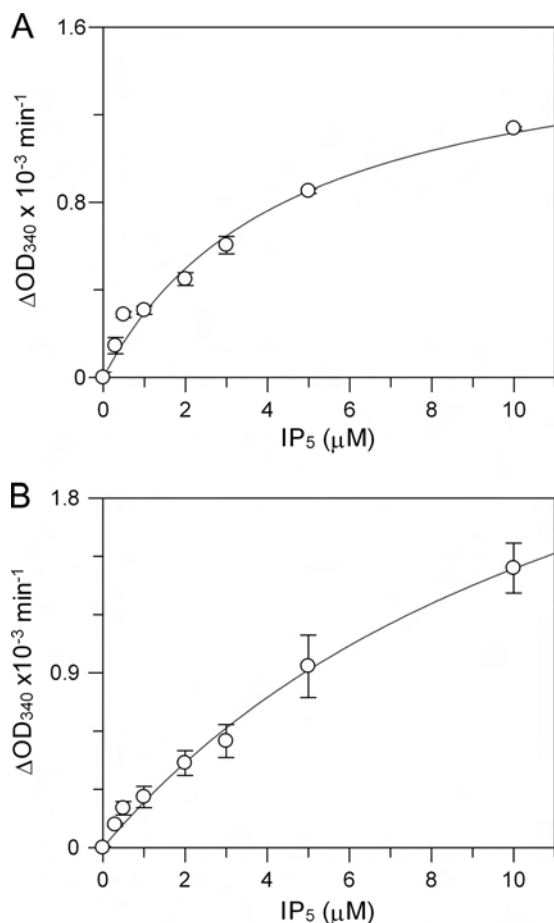
FIGURE 3. Snapshots of the different IP<sub>5</sub> 2-K conformations. A, sequential view of IP<sub>5</sub> 2-K active site structure in open (red), half-closed (orange), and closed (blue) forms. The N- and C-lobes are distinguished by different color shades. The key protein elements that cover the active site and that change conformations between the different forms (G-loop, L3, α6) are highlighted in green. Interactions for clasp formation in the closed form as well as the singular intramolecular ion pair (Asp<sup>255</sup>–Lys<sup>200</sup>) in open and half-closed forms are highlighted. Sulfate moieties are shown as green sticks. B, IP<sub>5</sub> 2-K active site along its different forms represented as sticks and following the code colors of A. Specific interactions formed and broken upon nucleotide and inositide binding are shown. Sulfates and substrates are shown as black sticks, highlighting phosphates and sulfate atoms as blue spheres.

markedly the discrimination of the enzyme in favor of Ins(3,4,5,6)P<sub>4</sub>. These results highlight the importance of Arg<sup>130</sup> in P1 binding. We are not aware of other structure-based studies of discrimination between enantiomeric substrates in inositol phosphate kinases.

**Intrinsic Fluorescence Decreases upon Substrate Binding**—The new structures that we have obtained provide description of conformational changes consequent upon ligand binding. We have undertaken ligand binding analysis of native and W129A mutants by fluorescence. For this purpose, we titrated Ins(1,3,4,5,6)P<sub>5</sub> substrate and recorded changes in intrinsic tryptophan fluorescence, exciting at 295 nm to limit fluorescence to tryptophan residues. We performed similar experiments titrating the nucleotide analog AMPPNP. Before performing substrate/analog titrations, we undertook Stern-Volmer analysis of quenching of tryptophan fluorescence with iodide as the quencher. The linearity of the unmodified Stern-Volmer plot and the y axis intercept of 1.03 for native and 1.02 for W129A (data not shown) indicate that the tryptophan residues of native and W129A protein are all accessible to the quenching agent. We also assessed the temperature dependence of fluorescence quenching by iodide. The increased gradient and linearity of the plot at higher temperature indicate that the

quenching process is collisional. These data are compatible with all tryptophans solvent-exposed in native and W129A protein. IP<sub>5</sub> 2-K has a total of four tryptophan residues among which two of them are on protein surface but partially buried in an hydrophobic pocket (Trp<sup>13</sup> and Trp<sup>69</sup>), although the other two are totally exposed (Trp<sup>129</sup> and Trp<sup>381</sup>). More precisely, Trp<sup>129</sup> would change its conformation from a partially exposed (closed form) to a completely exposed (open form, mutated by Ala) upon binding.

Incremental addition of Ins(1,3,4,5,6)P<sub>5</sub> yielded a saturating decrease in fluorescence for both native and W129A IP<sub>5</sub> 2-K (Fig. 6, A and B). Although it is clear that the decrease in fluorescence for the W129A mutant arises from residues other than tryptophan Trp<sup>129</sup>, that for native protein additionally includes Trp<sup>129</sup>. Fluorescence decreases in both cases are moderate. Assuming that Trp<sup>281</sup>, permanently solvent-exposed, is not a main contributor, the moderate changes are in concordance with slight distortions produced in Trp<sup>13</sup> and Trp<sup>69</sup> pockets upon inositide binding observed in the crystals. Plots of fractional change in fluorescence versus ligand such as that shown for native protein (Fig. 6C) yielded *K<sub>d</sub>* values for Ins(1,3,4,5,6)P<sub>5</sub> of 0.35 ± 0.12 μM for native protein and 0.42 ± 0.03 μM for W129A IP<sub>5</sub> 2-K. These data show that binding of



**FIGURE 4. Kinetic analysis of IP<sub>5</sub> 2-K.** Enzyme activity was assayed by a coupled enzyme assay. Titration of Ins(1,3,4,5,6)P<sub>5</sub> yielded saturating increase in enzyme activity of native (A) and W129A IP<sub>5</sub> 2-K (B). Progress of reaction curves were fitted by nonlinear regression to the Michaelis-Menten equation. The data in the text provide the mean  $\pm$  S.D. of kinetic parameters determined in three independent experiments with 4 to 5 replicate samples.

**TABLE 3**

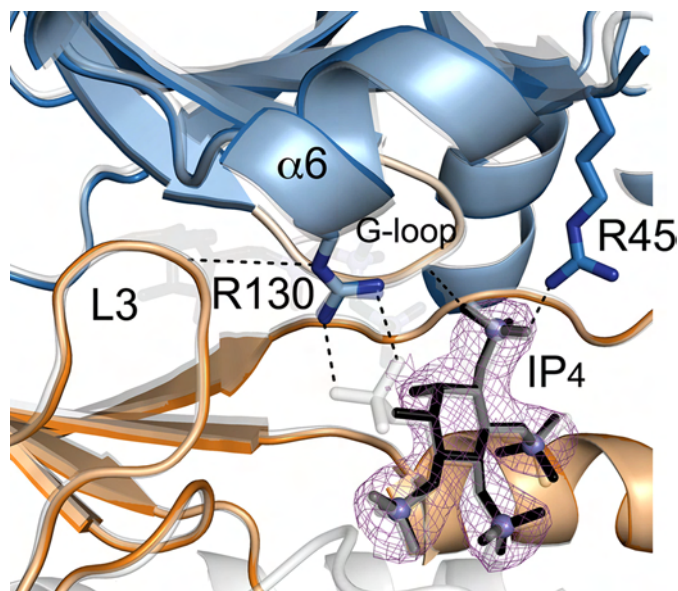
**Kinetic parameters of wild type and W129A IP<sub>5</sub> 2-K**

Kinetic parameters were determined by nonlinear least squares fitting to the Michaelis-Menten equation of plots of reaction velocity *versus* Ins(1,3,4,5,6)P<sub>5</sub> concentration, mean  $\pm$  S.D. of three experiments with 4 to 5 replicates per experiment.

	$V_{\max \text{ app}}$ $\mu\text{mol/min/mg}$	$K_m \text{ app}$ $\mu\text{M}$
Wild type	0.54 (0.07)	3.7 (1.0)
W129A	0.95 (0.16)	13.7 (3.0)

Ins(1,3,4,5,6)P<sub>5</sub> causes conformational change in IP<sub>5</sub> 2-K (either wild type or mutant), a result consistent with differences in conformation found between ligand-free and IP<sub>5</sub> 2-K inositide-bound structures. Remarkably, the affinity for inositide is similar in both proteins.

Titration of AMPPNP also gave a saturating decrease in fluorescence, again for both native and W129A IP<sub>5</sub> 2-K (Fig. 6, D and E). Plots of fractional change in fluorescence *versus* ligand and such as that shown for native protein (Fig. 6F) yielded  $K_d$   $37 \pm 9 \mu\text{M}$  and  $55 \pm 2 \mu\text{M}$ , respectively, for native and W129A protein. These structural changes reported by fluorescence upon AMPPNP binding are in agreement with the crystallographic differences in conformation between W129A IP<sub>5</sub> 2-K ligand-free and in complex with the nucleotide. The affinities for the



**FIGURE 5. IP<sub>5</sub> 2-K binary complex with AMPPNP and Ins(3,4,5,6)P<sub>4</sub>.** The N-lobe is shown in blue, the C-lobe in orange, and IP<sub>4</sub> as black sticks with phosphates highlighted as slate spheres (AMPPNP is shown as transparent sticks). The superposed IP<sub>5</sub> 2-K-AMPPNP-IP<sub>5</sub> complex (Protein Data Bank code 2xan) is shown as transparent white schematics and sticks. The inositide ligands from the N-lobe are shown as blue sticks. All protein interactions with inositide are retained with the exception of Arg<sup>130</sup>-P1 formed with IP<sub>5</sub>. However the position of Arg<sup>130</sup> is the same, and the interaction with L3 is conserved in this complex.

**TABLE 4**

**Relative activities of IP<sub>5</sub> 2-K mutants for enantiomeric Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub> substrates**

Activities of mutants are expressed relative to the activity of wild-type enzyme assayed on the same microplate. Results obtained from initial rate measurements are also expressed as a ratio of enzyme activities against the enantiomeric substrates Ins(1,4,5,6)P<sub>4</sub>/Ins(3,4,5,6)P<sub>4</sub>. Assays were performed at 5  $\mu\text{M}$  IP<sub>4</sub>, 100  $\mu\text{M}$  ATP. The experiment has been repeated three times with similar results. A similar experiment with a different batch of substrate yielded Ins(1,4,5,6)P<sub>4</sub>/Ins(3,4,5,6)P<sub>4</sub> ratios for WT, 0.64  $\pm$  0.04; W129A, 0.58  $\pm$  0.03, and W129V/R130A, 0.10  $\pm$  0.05.

	Ins(1,4,5,6)P <sub>4</sub>		Ins(3,4,5,6)P <sub>4</sub>		Ratio	
	% of WT	S.D.	% of WT	S.D.	Mean	S.D.
Wild type	100		100		0.99	0.15
R45A	179	23	135	8	1.31	0.17
W129A	75	7	97	15	0.77	0.12
R130A	17	7	59	5	0.28	0.07

nucleotide showed by the wild type and W129A IP<sub>5</sub> 2-K are quite similar, although slightly decreased for the W129A protein. Because the clasp would not be properly formed in this mutant, probably a less stable nucleotide complex is formed.

## DISCUSSION

We have obtained different snapshots of IP<sub>5</sub> 2-K functional states (Figs. 2A and 3) by x-ray crystallography. Each different state is unequivocally correlated with binding of nucleotide or inositide at the active site. Thus, the protein shows an open conformation in the absence of substrates, a half-closed conformation when bound to nucleotide, and as reported previously (24), a closed conformation in the inositide-bound state. In agreement with the crystal structures presented here, IP<sub>5</sub> 2-K intrinsic fluorescence decreases upon nucleotide and inositide binding (Fig. 6). We have predicted previously such conformational changes (24). Very recently, some experimental evidence of these changes has been reported by partial proteolysis exper-



## IP<sub>5</sub> 2-K Conformational Changes upon Substrate Binding

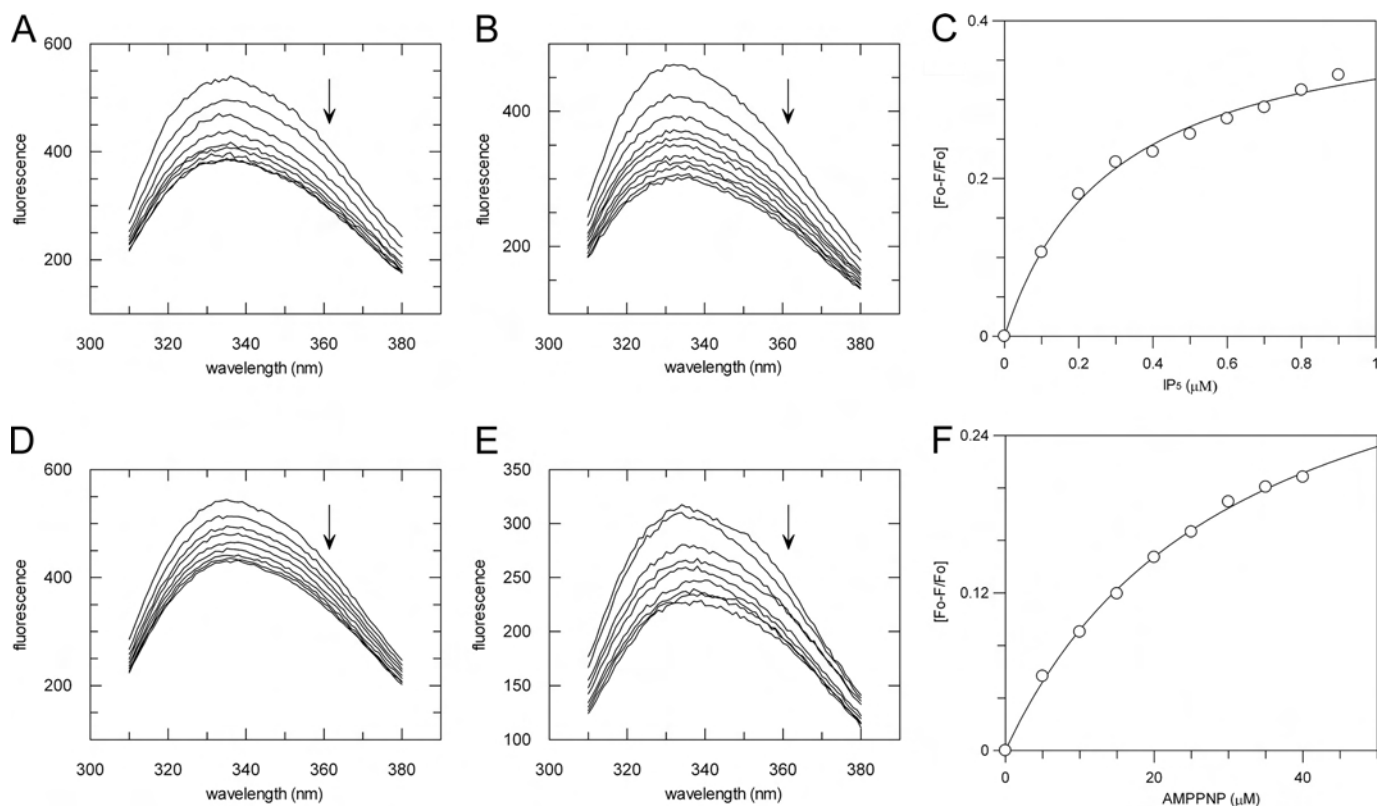


FIGURE 6. **Ligand binding to IP<sub>5</sub> 2-K.** Titration of Ins(1,3,4,5,6)P<sub>5</sub> (A and B) or AMPPNP (D and E) yielded saturating decreases in intrinsic fluorescence of native (A and D) and W129A IP<sub>5</sub> 2-K (B and E). Arrows indicate the direction of change of fluorescence on incremental additions of ligand. Plots of fractional change in fluorescence are shown for native protein versus Ins(1,3,4,5,6)P<sub>5</sub> (C) and versus AMPPNP (F).

iments that show stabilization of some regions of IP<sub>5</sub> 2-K upon inositide and nucleotide binding (40). This work shows that high resolution analysis of the conformational events has been possible by a single mutation (W129A) allowing description of a complete picture of the different conformational states experienced by IP<sub>5</sub> 2-K throughout the substrate binding stages of catalysis. The assumption that the W129A mutant is a mimic of WT IP<sub>5</sub> 2-K is based on the experimental evidence presented here. In addition, the W129A mutant presents similar protection from proteolysis in the presence of nucleotide and inositide as wild type IP<sub>5</sub> 2-K, as well as a similar band shift on native gels in presence of IP<sub>6</sub> (data not shown).

**Inositide Alone Can Promote Closed Conformation and Clasp Formation**—An intriguing aspect from our studies is that inositide is sufficient and necessary to trigger the closed conformation and complete the clasp formation, and it can do so independently of ATP binding. We were able to crystallize IP<sub>5</sub> 2-K with AMPPNP/Ins(3,4,5,6)P<sub>4</sub> in the closed form; however, trials with AMPPNP/Ins(1,4,5,6)P<sub>4</sub> were unsuccessful. One potential explanation of these observations is that the inositide P3 coordination to Arg<sup>45</sup> and/or the G-loop is essential to favor the closed form upon inositide binding (Fig. 5). In fact, because Arg<sup>45</sup> mutation does not significantly compromise enzyme activity, we suggest that the G-loop may be the principal N-lobe element for inositide binding and subsequent conformational changes. In addition, the clasp is fully formed only in the closed conformation. Mutation of clasp elements that would preclude its formation yield active enzymes, suggesting that the clasp formation is not essential for protein activity (Table 2). On the

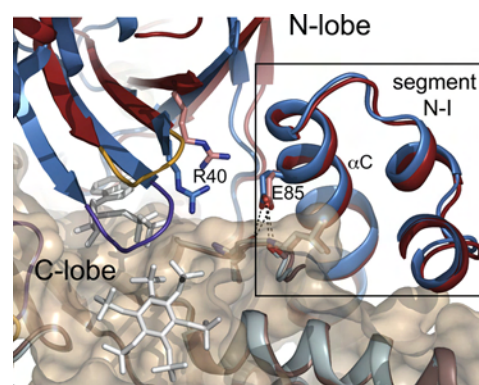
contrary, the clasp might be relevant for modulating protein kinetics as suggested by enzymatic and kinetic measurements. In conclusion, both events, domain closure and clasp formation, are coupled to inositide binding, and the G-loop could be playing a direct role in the conformational changes. A main role for a G-loop has been largely studied in protein kinases, where it has been suggested that it is a regulatory flap above the ATP-binding site showing great flexibility (41, 42).

In this closed form, the inositide blocks the nucleotide entrance/exit to/from the active site. This suggests that the nucleotide might bind first, producing a half-closed conformation, and the inositide later completed the domain closure. Then the product release should proceed in an inverse order, because the nucleotide is completely occluded in the inositide-bound form. The exit of ADP could be the limiting step for IP<sub>5</sub> 2-K activity, as it has been reported for PKs (43), because the motion to open the domains and/or to undo the clasp is necessary for nucleotide release. These suggestions are supported by the fact that clasp-affected mutants as W129A or ΔS253/ΔE255 produce more efficient or active proteins. Nevertheless, more work has to be done to confirm these entire hypotheses about substrate binding order. We consider that the capacity of inositide alone to trigger the closed conformation could reflect its role as an enzyme inhibitor and therefore in regulation of enzymatic activity. Several examples of similar proteins being inhibited by its substrate/product can be found in the literature (44–47).

**Modulation of Enantiomeric IP<sub>4</sub> Preference of Plant IP<sub>5</sub> 2-Ks**—The relative contribution of different pathways to IP<sub>6</sub> synthesis in plants is the focus of much attention (18, 19). We

have compared the activity of IP<sub>5</sub> 2-K toward Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub> for native protein and for a range of mutants, because both enantiomers represent intermediates in disparate pathways reported to proceed from inositol 1,4,5-trisphosphate and inositol, respectively (19). Ins(1,4,5,6)P<sub>4</sub> is a product of plant multikinase action against inositol 1,4,5-trisphosphate, certainly *in vitro* (48, 49). *AtIPK2β*, a multikinase, and *AtIPK1* (IP<sub>5</sub> 2-K) contribute to IP<sub>6</sub> synthesis in *Arabidopsis* (18). Ins(3,4,5,6)P<sub>4</sub> was identified in early studies of IP<sub>6</sub> synthesis in duckweed (50). This isomer is a substrate of plant ATP-grasp fold inositol tris/tetrakisphosphate kinases (20, 51–53), and enzymes of this class contribute to IP<sub>6</sub> synthesis in maize (20). We have shown using a coupled assay that Ins(3,4,5,6)P<sub>4</sub> and Ins(1,4,5,6)P<sub>4</sub> are equally effective *in vitro* substrates for IP<sub>5</sub> 2-K, and we have now determined the role of Arg<sup>130</sup> and Arg<sup>45</sup> in the specification of substrate preference between enantiomers. We note that although Arg<sup>130</sup> is fully conserved across kingdoms, Arg<sup>45</sup> is not, and within plants the equivalent residue is variable (6). Moreover, despite the elucidation of structures of a variety of inositol phosphate kinases belonging to the ATP-grasp fold and inositol polyphosphate kinase families, this is the first study in which discrimination between enantiomeric substrates has been investigated. Our choice of substrates, although dictated by consideration of what is known of inositol phosphate metabolism in plants, is of obvious relevance to animal inositol phosphate metabolism centered on the same enantiomers. In addition, manipulation or inverting the enantiomeric substrate preference and therefore enantiomeric product production is of great interest from a chemical and a biotechnological point of view.

**Singular Enzyme among the Inositide Kinase Family**—The structural features for all inositide kinase families have been extensively studied to determine the structural determinants of substrate specificity. Structures are available for different inositol phosphate kinases in the presence/absence of substrates and products, providing data about structural changes produced upon binding of nucleotide and/or inositide. With respect to the IPK family, to which IP<sub>5</sub> 2-K belongs, none of the structures solved up to now (IP<sub>3</sub> 3-K (21, 22) and IPmK (23)) show significant structural changes upon nucleotide binding. However, data upon inositide binding is only known for IP<sub>3</sub> 3-K, for which a slight conformational change is described. This change consists of an inositide site rearrangement achieved by a protein intramolecular bridge breakage between the C-lobe and IP-lobe (Arg<sup>419</sup>–Asp<sup>374</sup>) leaving Arg<sup>419</sup> free to interact with inositide. As seen in IP<sub>5</sub> 2-K, a bond is also broken upon inositide binding (Lys<sup>200</sup>–Asp<sup>255</sup>) to leave Lys<sup>200</sup> able to bind inositide. Nevertheless, the bridges in the two proteins are not structurally related, and IP<sub>3</sub> 3-K does not present structurally equivalent elements of L3 or α6. In any case, the conformational change produced in IP<sub>3</sub> 3-K is local, and it does not involve large movements between domains as in IP<sub>5</sub> 2-K. A member from the other family of inositol phosphate kinases, those from ATP-grasp fold, also shows structural differences upon inositide binding. In the recent published structure for the pyrophosphate inositol-synthesizing enzyme PPIP5K (26), the side chain of three large basic residues changes markedly upon inositide binding, revealing a major reorganization in the



**FIGURE 7. IP<sub>5</sub> 2-K αC and C-lobe are coupled during domain motions.** The IP<sub>5</sub> 2-K N-lobes are shown as red (open form) and blue (closed form) schematics, respectively. The C-lobe is shown as surface representation. Element N-I αC is *squared*. Glu<sup>85</sup> does not form a bridge with Arg<sup>40</sup>, unlike the situation in PKs in active conformation. Glu<sup>85</sup> maintains interaction with the C-lobe through the motion. The residues involved are shown as sticks.

inositide-binding site configuration without involving backbone changes. In conclusion, this is the first time that such a conformational change involving domain movements upon inositide as well as nucleotide binding is shown for a member of the inositol phosphate kinase family. However, we must not ignore the fact that it has not yet been possible to crystallize every inositol kinase and in all their different states. Thus, none of the IP<sub>6</sub> K isoforms structures are available at the moment and neither are the IPmKs in the presence of inositide nor PPIP5K in the absence of nucleotide. Consequently, the possibility of undetermined conformational changes for these families of enzymes cannot be precluded.

**Comparison with Protein Kinases**—A similar domain movement has been reported for protein kinases (42, 54). In the most studied PK (PKA), the nucleotide binding generates a closure between both N- and C-lobes (43). In our previous work (24), we reported that the distinctive bridge linking αC and the N-lobe β3 in PKAs structures representing the active conformation (Lys<sup>72</sup>–Glu<sup>91</sup>) is not formed in IP<sub>5</sub> 2-K (Arg<sup>40</sup>–Glu<sup>85</sup> equivalent residues); therefore, we suggested that the bridge is not a hallmark for active conformation in IP<sub>5</sub> 2-Ks (Fig. 7). On the contrary, interaction between αC and the C-lobe is retained in both PKAs and IP<sub>5</sub> 2-Ks. Thus, in concordance with our previous results, the IP<sub>5</sub> 2-K open and half-closed conformations also lack this bridge, and in the domain motion reported here, the N-lobe region containing αC (N-I) is coupled with the C-lobe (Fig. 7), whereas in PKA the motion from open to close conformations keeps αC in contact with both lobes (43).

**Conclusions**—From our previous inositide-bound IP<sub>5</sub> 2-K structure (closed form) and using a rational approach, we have designed several IP<sub>5</sub> 2-K mutations prone to yield an open protein conformation. From a specific protein mutant, W129A, we have conducted crystallization of IP<sub>5</sub> 2-K in an open, free of substrates, and a half-closed nucleotide-bound conformation. Combining this information with the known closed inositide-bound form, we have provided a schematic view of the conformations adopted by IP<sub>5</sub> 2-K in the course of IP<sub>6</sub> synthesis. This is the first time that such a conformational change is reported for an inositide kinase. Moreover, these changes explain all the experimental observations, including fluorescence measurements, that pointed to a nucleo-

tide- and inositide-induced conformational changes in IP<sub>5</sub> 2-K. The different protein forms captured can be considered as a “biological mimic state,” proving mutagenesis to be a successful tool in Protein Crystallography.

A general problem facing the protein kinase field is the design of selective inhibitors, because all protein kinases bind ATP in a very similar manner. In the case of inositol phosphate kinases, we can use the inositide-binding site for design of specific inhibitors; however, a number of crucial inositide kinases could be affected by similar molecules. To overcome this selectivity problem and avoid off-target effects, an understanding of the structural basis of substrate specificity and preference is essential. Furthermore, a common approach is to exploit protein conformational changes to design specific inhibitors. In this context, docking studies, which start from an ensemble of different states, have been proven to be successful. The main goals of this work were to provide new substrate recognition features, as elements involved in enantiotopic phosphate discrimination and structures of different protein conformational states. These are valuable tools in the search for specific and selective protein inhibitors with potential applications as biological tools for *in vivo* studies. The implications of IP<sub>5</sub> 2-K in crucial cellular and developmental events have been already proven; however, elaboration of protein inhibitors could help in the determination of further roles for IP<sub>5</sub> 2-K in mammals.

**Acknowledgments**—We thank María Álvarez-Cao, Juana María González-Rubio, and Kendall Baker for technical assistance and the European Synchrotron Radiation Facility (Grenoble, France) for providing beam time and assistance during data collection. We thank Stephen Mills (University of Bath, UK) for the gift of Ins(1,4,5,6)P<sub>4</sub> and Ins(3,4,5,6)P<sub>4</sub> used in some of the enzyme assays described.

## REFERENCES

1. Streb, H., Irvine, R. F., Berridge, M. J., and Schulz, I. (1983) Release of Ca<sup>2+</sup> from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol 1,4,5-trisphosphate. *Nature* **306**, 67–69
2. Irvine, R. F., and Schell, M. J. (2001) Back in the water. The return of the inositol phosphates. *Nat. Rev. Mol. Cell Biol.* **2**, 327–338
3. Ives, E. B., Nichols, J., Wente, S. R., and York, J. D. (2000) Biochemical and functional characterization of inositol 1,3,4,5,6-pentakisphosphate 2-kinases. *J. Biol. Chem.* **275**, 36575–36583
4. Phillippy, B. Q., Ullah, A. H., and Ehrlich, K. C. (1994) Purification and some properties of inositol 1,3,4,5,6-pentakisphosphate 2-kinase from immature soybean seeds. *J. Biol. Chem.* **269**, 28393–28399
5. Stephens, L. R., Hawkins, P. T., Stanley, A. F., Moore, T., Poyner, D. R., Morris, P. J., Hanley, M. R., Kay, R. R., and Irvine, R. F. (1991) Myo-inositol pentakisphosphates. Structure, biological occurrence, and phosphorylation to myo-inositol hexakisphosphate. *Biochem. J.* **275**, 485–499
6. Sun, Y., Thompson, M., Lin, G., Butler, H., Gao, Z., Thornburgh, S., Yau, K., Smith, D. A., and Shukla, V. K. (2007) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize. Molecular and biochemical characterization. *Plant Physiol.* **144**, 1278–1291
7. Sweetman, D., Johnson, S., Caddick, S. E., Hanke, D. E., and Brearley, C. A. (2006) Characterization of an *Arabidopsis* inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). *Biochem. J.* **394**, 95–103
8. Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W., and Wente, S. R. (2002) The synthesis of inositol hexakisphosphate. Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *J. Biol. Chem.* **277**, 31857–31862
9. Verbsky, J., Lavine, K., and Majerus, P. W. (2005) Disruption of the mouse inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8448–8453
10. Alcázar-Román, A. R., Bolger, T. A., and Wente, S. R. (2010) Control of mRNA export and translation termination by inositol hexakisphosphate requires specific interaction with Gle1. *J. Biol. Chem.* **285**, 16683–16692
11. Shen, X., Xiao, H., Ranallo, R., Wu, W. H., and Wu, C. (2003) Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* **299**, 112–114
12. Sarmah, B., Latimer, A. J., Appel, B., and Wente, S. R. (2005) Inositol polyphosphates regulate zebrafish left-right asymmetry. *Dev. Cell* **9**, 133–145
13. Agarwal, R., Mumtaz, H., and Ali, N. (2009) Role of inositol polyphosphates in programmed cell death. *Mol. Cell. Biochem.* **328**, 155–165
14. Bennett, M., Onnebo, S. M., Azevedo, C., and Saiardi, A. (2006) Inositol pyrophosphates. Metabolism and signaling. *Cell. Mol. Life Sci.* **63**, 552–564
15. Murphy, A. M., Otto, B., Brearley, C. A., Carr, J. P., and Hanke, D. E. (2008) A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *Plant J.* **56**, 638–652
16. Raboy, V. (2001) Seeds for a better future. “Low phytate” grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci.* **6**, 458–462
17. Raboy, V. (2003) myo-Inositol-1,2,3,4,5,6-hexakisphosphate. *Phytochemistry* **64**, 1033–1043
18. Stevenson-Paulik, J., Bastidas, R. J., Chiou, S. T., Frye, R. A., and York, J. D. (2005) Generation of phytate-free seeds in *Arabidopsis* through disruption of inositol polyphosphate kinases. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12612–12617
19. Raboy, V. (2007) The ABCs of low-phytate crops. *Nat. Biotechnol.* **25**, 874–875
20. Shi, J., Wang, H., Wu, Y., Hazebroek, J., Meeley, R. B., and Ertl, D. S. (2003) The maize low phytic acid mutant *lpa2* is caused by mutation in an inositol phosphate kinase gene. *Plant Physiol.* **131**, 507–515
21. González, B., Schell, M. J., Letcher, A. J., Veprintsev, D. B., Irvine, R. F., and Williams, R. L. (2004) Structure of a human inositol 1,4,5-trisphosphate 3-kinase. Substrate binding reveals why it is not a phosphoinositide 3-kinase. *Mol. Cell* **15**, 689–701
22. Miller, G. J., and Hurley, J. H. (2004) Crystal structure of the catalytic core of inositol 1,4,5-trisphosphate 3-kinase. *Mol. Cell* **15**, 703–711
23. Holmes, W., and Jögl, G. (2006) Crystal structure of inositol phosphate multikinase 2 and implications for substrate specificity. *J. Biol. Chem.* **281**, 38109–38116
24. González, B., Baños-Sanz, J. I., Villate, M., Brearley, C. A., and Sanz-Aparicio, J. (2010) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide-binding site for axial 2-OH recognition. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9608–9613
25. Miller, G. J., Wilson, M. P., Majerus, P. W., and Hurley, J. H. (2005) Specificity determinants in inositol polyphosphate synthesis. Crystal structure of inositol 1,3,4-trisphosphate 5/6-kinase. *Mol. Cell* **18**, 201–212
26. Wang, H., Falck, J. R., Hall, T. M., and Shears, S. B. (2012) Structural basis for an inositol pyrophosphate kinase surmounting phosphate crowding. *Nat. Chem. Biol.* **8**, 111–116
27. Baños-Sanz, J. I., Villate, M., Sanz-Aparicio, J., Brearley, C. A., and González, B. (2010) Crystallization and preliminary x-ray diffraction analysis of inositol 1,3,4,5,6-pentakisphosphate kinase from *Arabidopsis thaliana*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **66**, 102–106
28. Baños-Sanz, J. I., Sanz-Aparicio, J., Brearley, C. A., and González, B. (2012) Expression, purification, crystallization, and preliminary x-ray diffraction analysis of the apo-form of IP<sub>5</sub>-2K from *Arabidopsis thaliana*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **68**, 701–704
29. Batty, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281
30. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82
31. Collaborative Computational Project Number 4 (1994) The CCP4 suite. Programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*



- 50, 760–763
32. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
33. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
34. Emsley, P., and Cowtan, K. (2004) Coot. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
35. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK. A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291
36. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*. DeLano Scientific LLC, San Carlos, CA
37. Hayward, S., and Lee, R. A. (2002) Improvements in the analysis of domain motions in proteins from conformational change. DynDom Version 1.50. *J. Mol. Graph. Model.* **21**, 181–183
38. Mayr, G. W., Windhorst, S., and Hillemeier, K. (2005) Antiproliferative plant and synthetic polyphenolics are specific inhibitors of vertebrate inositol-1,4,5-trisphosphate 3-kinases and inositol polyphosphate multikinase. *J. Biol. Chem.* **280**, 13229–13240
39. Rowan, A. S., Nicely, N. I., Cochrane, N., Wlassoff, W. A., Claiborne, A., and Hamilton, C. J. (2009) Nucleoside triphosphate mimicry. A sugar triazolyl nucleoside as an ATP-competitive inhibitor of B. anthracis pantothenate kinase. *Org. Biomol. Chem.* **7**, 4029–4036
40. Gosein, V., Leung, T. F., Krajden, O., and Miller, G. J. (2012) Inositol phosphate-induced stabilization of inositol 1,3,4,5,6-pentakisphosphate 2-kinase and its role in substrate specificity. *Protein Sci.* **21**, 737–742
41. Bossemeyer, D. (1994) The glycine-rich sequence of protein kinases. A multifunctional element. *Trends Biochem. Sci.* **19**, 201–205
42. Taylor, S. S., Radzio-Andzelm, E., Madhusudan, Cheng, X., Ten Eyck, L., and Narayana, N. (1999) Catalytic subunit of cyclic AMP-dependent protein kinase. Structure and dynamics of the active site cleft. *Pharmacol. Ther.* **82**, 133–141
43. Cox, S., Radzio-Andzelm, E., and Taylor, S. S. (1994) Domain movements in protein kinases. *Curr. Opin. Struct. Biol.* **4**, 893–901
44. Abdullah, M., Hughes, P. J., Craxton, A., Gigg, R., Desai, T., Marecek, J. F., Prestwich, G. D., and Shears, S. B. (1992) Purification and characterization of inositol-1,3,4-trisphosphate 5/6-kinase from rat liver using an inositol hexakisphosphate affinity column. *J. Biol. Chem.* **267**, 22340–22345
45. Nalaskowski, M. M., Bertsch, U., Fanick, W., Stockebrand, M. C., Schmale, H., and Mayr, G. W. (2003) Rat inositol 1,4,5-trisphosphate 3-kinase C is enzymatically specialized for basal cellular inositol trisphosphate phosphorylation and shuttles actively between nucleus and cytoplasm. *J. Biol. Chem.* **278**, 19765–19776
46. Voglmaier, S. M., Bembenek, M. E., Kaplin, A. I., Dormán, G., Olszewski, J. D., Prestwich, G. D., and Snyder, S. H. (1996) Purified inositol hexakisphosphate kinase is an ATP synthase. Diphosphoinositol pentakisphosphate as a high energy phosphate donor. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 4305–4310
47. Van Rooijen, L. A., Rossowska, M., and Bazan, N. G. (1985) Inhibition of phosphatidylinositol-4-phosphate kinase by its product phosphatidylinositol 4,5-bisphosphate. *Biochem. Biophys. Res. Commun.* **126**, 150–155
48. Stevenson-Paulik, J., Odom, A. R., and York, J. D. (2002) Molecular and biochemical characterization of two plant inositol polyphosphate 6-/3-/5-kinases. *J. Biol. Chem.* **277**, 42711–42718
49. Xia, H. J., Brearley, C., Elge, S., Kaplan, B., Fromm, H., and Mueller-Roeber, B. (2003) *Arabidopsis* inositol polyphosphate 6-/3-kinase is a nuclear protein that complements a yeast mutant lacking a functional ArgR-Mcm1 transcription complex. *Plant Cell* **15**, 449–463
50. Brearley, C. A., and Hanke, D. E. (1996) Inositol phosphates in the duckweed *Spirodela polyrrhiza* L. *Biochem. J.* **314**, 215–225
51. Sweetman, D., Stavridou, I., Johnson, S., Green, P., Caddick, S. E., and Brearley, C. A. (2007) *Arabidopsis thaliana* inositol 1,3,4-trisphosphate 5/6-kinase 4 (*AtITPK4*) is an outlier to a family of ATP-grasp fold proteins from *Arabidopsis*. *FEBS Lett.* **581**, 4165–4171
52. Josefsen, L., Bohn, L., Sørensen, M. B., and Rasmussen, S. K. (2007) Characterization of a multifunctional inositol phosphate kinase from rice and barley belonging to the ATP-grasp superfamily. *Gene* **397**, 114–125
53. Caddick, S. E., Harrison, C. J., Stavridou, I., Mitchell, J. L., Hemmings, A. M., and Brearley, C. A. (2008) A *Solanum tuberosum* inositol phosphate kinase (StITPK1) displaying inositol phosphate-inositol phosphate and inositol phosphate-ADP phosphotransferase activities. *FEBS Lett.* **582**, 1731–1737
54. Hyeon, C., Jennings, P. A., Adams, J. A., and Onuchic, J. N. (2009) Ligand-induced global transitions in the catalytic domain of protein kinase A. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3023–3028



